Measurement and Prediction of Protein Phase Behaviour and Protein-Protein Interactions

Faber, Cornelius

Publication date:
2006

Document Version
Publisher's PDF, also known as Version of record

Measurement and Prediction of Protein Phase Behaviour and Protein-Protein-Interactions

---

by

Cornelius Faber

Ph.D. Thesis

Center for Microbial Biotechnology
Biocentrum-DTU
Technical University of Denmark

March 2006

Preface

This thesis is the result of my studies for the Ph.D.-degree performed at the Center for Microbial Biotechnology, Biocentrum-DTU, Technical University of Denmark and Novozymes A/S, Bagsværd, Denmark in the period from February 2003 to January 2006.

First of foremost I wish to thank my supervisors Dr. Svend Kaasgaard, Associate Professor Tim Hobley, Professor Jørgen Mollerup and Professor Owen Thomas for providing indispensable scientific and practical support and constructive feedback throughout this study.

During this study I could supervise Louise Lyhne, Rune Prior and Mikkel Nielsen during their master's projects. I very much appreciate their high motivation during their projects and I believe that the cooperation was beneficial for them as well as for me.

I want to thank all the members of the Downstream Processing Group at DTU as well as the employees of the Liquid Product Development department at Novozymes who provided a pleasant working environment. In particular, I want to emphasise the never-failing support from Hanne Margrethe Nielsen, Ewy Hein Møller and Kim Bruno Andersen.

Parts of this work have been conducted in external labs. I wish to thank André Dumetz and Professor Abraham M. Lenhoff from the Department of Chemical Engineering, University of Delaware, USA, for having provided the opportunity to enrich my scientific work by many precious aspects. Furthermore, I wish to thank Dr. Fabienne Espitalier from the Laboratoire de Génie des Procédés des Solides Divisés, Ecole des Mines d'Albi-Carmaux in France for having paved the way for a short but fruitful stay in her department.

Last but not least I wish to thank the Novozymes Bioprocess Academy for funding my Ph.D.-scholarship.

Cornelius Faber, March 2006

Dansk Resumé

Hovedformålet med denne Ph.D. afhandling var at evaluere traditionelle og nye indgangsvinkler til karakterisering og kvantificering af proteiners opløselighed og krystalvækst under indflydelse af opløsningens egenskaber som f. eks. forandringer i salt koncentrationen, pH og temperaturforskelle. Der blev desuden lavet overordnede fasediagrammer under udvalgte betingelser. Metoderne blev evalueret med det formål at reducere protein forbruget, også når der arbejdes med mindre rene enzymer. Til den eksperimentelle del af denne afhandling blev der brugt enzymkoncentrater af to rekombinante α-amylaser fra *Bacillus halmapalus* (BHA) og *Bacillus licheniformis* (BLA), som var blevet oprenset til teknisk grad.

Udviklingen af optimerede mikrobielle stammer sammen med fremskreden gæringsteknologi fører til et renere industrielt gæringsmedie med højere produkt koncentration, som vil muliggøre en økonomisk masse-produktion af proteiner af lav værdi. Samtidig øges risikoen betragteligt for at operere over opløslighedsgrænsen og deraf følgende ukontrolleret precipitering eller krystalformation, hvilket kan føre til væsentlige problemer i produktionsforløbet, f. eks. produkttab under almindelige enhedsoperations til fjernelse af biomasse, som f. eks. centrifugering eller filtrering. Udvikling og gennemførelse af downstream processer under betingelser, der ligger tæt på opløslighedsgrænsen er særlig udfordrende og afhænger ofte af trial-and-error pga. mangel på omfattende data for proteinopløslighed for processrelevante betingelser. Således er information omkring opløslighedsegenskaber af vidtrækkende værdi, ikke kun under udviklingen af rensningsprocesserne, men også ved den indledende screening rutiner af nye lovende molekyler. Opløslighedsdata er kun tilgængelig for et begrænset antal proteiner, eftersom traditionelle metoder for udarbejdelse af opløslighedskurver kræver store mængder af protein, samt at krystallisering af disse proteiner er mulig indenfor en rimelig tidsfrist – to kriterier som ofte ikke er til stede (kapitel 1).

Der er lavet en omfattende beskrivelse af BHA's oopløselighedsprofil (kapitel 2). Opløslighedskurverne blev opstillet ved hjælp af klassiske temperatur-kontrollerede batch krystalliserings processer i Eppendorfrør med et arbejdsvolumen på 1 mL. Indflydelsen af temperatur, pH og udvalgte kationer og anioner fra Hofmeister serien på opløsligheden blev

kvantificeret. Den viste sig at være stort set uafhængig af temperatur men stærkt afhængig af pH. Hofmeister serien for anioner blev fulgt i den korrekte rækkefølge, ligeledes blev rækken fulgt for de monovalente kationer, med undtagelse af lithium. Det var forventet at den ville være den ringeste precipitant, men viste sig at være den bedste. Der blev gennemført målinger af zeta-potentialet, og det viste at lithium øgede det isoelektriske punkt (pI) af BHA, hvilket kunne forklare den uventede effekt på enzymets opløslighed under tilstedeværelse af denne ion.

I kapitel 3 blev batch krystalliserings processer af BHA med udgangspunkt i tre forskellige overmættelser studeret nærmere. Krystallernes størrelsesfordeling, deres antal samt koncentrationen af opløst protein blev målt som funktion af tid. Allerede under pH justeringen for at skabe den overmættede opløsning fremkom en betragtelig mængde krystaller. Indenfor de første to til tre timer af processen øgedes krystal koncentrationen yderligere, men den forblev derefter konstant i en bestemt periode, hvis længde blev bestemt af graden af overmætning. Kun hvis overmætningen var stor nok, ville krystalkoncentrationen øges yderligere og en ligevægt opnået indefor den eksperimentelle tidshorizont på højest 48 timer. I modsætning til krystal koncentrationen ændrede fordelingen af krystal størrelserne sig ikke under processen. Den gennemsnitlige diameter af krystallerne forblev konstant og var mellem 4 og 5 µm. Således kunne man ikke konstatere fortsat vækst. Fordelingen af krystal størrelserne viste sig grundlæggende at være uafhængig af overmætningen.

Indflydelsen af udvalgte kationer og anioner fra Hofmeister serien på opløslighed af BLA er præsenteret i kapitel 4. I modsætning til BHA, er BLA stabil på begge sider af pI, hvilket muliggør at teste hypotesen om at Hofmeister serien bliver vendes om afhængigt af fortegnet på proteinets ladning. Opløsligheden af BLA blev målt ved pH 6, 7 og 8 ved salt koncentrationer på 0 M, 0,1 M og 0,5 M. Fortegnet på proteinets ladning i saltopløsningerne blev bestemt ved at måle zeta-potentialet ved de førnævnte pH-værdier, men kun i opløsningerne med 0 M og 0,1 M salt. Målinger af zeta-potentialet var upræcist i saltkoncentrationer på over 0,1 M, derfor kunne den resulterende pI ved 0,5 M ikke bestemmes. Ved alle testede betingelser blev den laveste opløslighed fundet ved pH 7, med undtagelse af 0,5 M natrium thiocyanate. En invertering af Hofmeister serien for kationer og anioner ved at ændre fortegnet på proteinets ladning kunne ikke bekræftes.

I kapitel 5 beskrives mikrotiterpladers potentiale for at frembringe BHA fasediagrammer bestående af zonerne precipitation og nukleation, såvel som de metastabile og undermættede zoner. Disse temperatur-kontrollerede mikro-batch eksperimenter blev gennemført ved en arbejdsvolumen på 200 µL og udfældningerne blev analyseret ved lysmikroskopi. Herved kunne betingelser for dannelse af amorfe udfældninger og forskellige typer krystaller identificeres som funktion af startbetingelserne. Efter 4 dages inkubation blev mikrotiterpladerne undersøgt, og sammen med opløseligheds data fra kapitel 2 kunne fasediagrammerne konstrueres. Disse blev bestemt som en funktion af pH, koncentrationen af natrium klorid ved pH 7 og 9, samt koncentrationen af natrium thiocyanat ved pH 7. I alle tilfælde kunne man identificere en bred nukleationszone. Amorft precipitat opstod kun med ved en høj initiel overmætning. For natrium klorid ved pH 9 blev der ikke fundet noget amorft precipitat. Den metastabile zone var meget smal for i systemer med natrium klorid ved pH 9, og som funktion af pH uden tilsat salt. Dog blev den metastabile zone bredere når natrium klorid koncentration øgedes, og den blev endnu bredere i tilstedeværelsen af natrium thiocyanat, begge dele blev målt ved pH 7.

Selvom proteinforbruget kunne reduceres betragteligt vha. de førnævnte metoder, så forbliver tilstedeværelsen af tilstrækkelige mængder af protein samt dets krystalliserbarhed en begrænsning som vil være svær at overvinde i traditionelle opløselighdes test. Derfor blev yderligere metoder undersøgt i kapitel 6, hvor BHA blev anvendt som model protein. Bestemmelsen af den anden osmotiske virial koefficient (B_{22}) er et attraktivt alternativ til målinger af proteiners opløslighed, og har den fordel at eksperimenter kan gennemføres i undermættede opløsninger, således at man ikke skal bruge store mængder protein og tidskrævende krystallisationsprocesser. B_{22} repræsenterer en Boltzmann-vægtet gennemsnitlig måling af protein-protein interaktioner, hvor positive B_{22}-værdier modsvarer frastødende interaktioner, og negative værdier modsvarer tiltrækkende interaktioner. Ved bestemmelsen af B_{22} kan man således forudsige processer for hhv. faseseparation og aggregering. Dynamisk lysspredning (DLS) af de her anvendte BHA preparationer viste tilstedeværelsen af signifikante mængder af opløslige enzymaggregater. Disse vil interferere ved bestemmelsen af B_{22} ved lysspredning, hvorfor det blev konkluderet at denne metode ikke var velegnet til bestemme B_{22} i BHA opløsninger. En alternativ metode til lysspredning er self-interaction chromatography (SIC), som før har vist sig at kunne tilvejebringe B_{22}-værdier, som kvantitativt stemmer overens med værdierne opnået vha. lysspredning, men som kan siges at være noget mere effektiv hvad angår forbrug af protein og tid. Eksperimenterne fra kapitel 6

bekræfter, at SIC metoden ikke påvirkes af aggregater. Derudover blev en stor mængde B_{22}-data genereret ved SIC, og der blev fundet en god overensstemmelse mellem B_{22} og opløsligheden af BHA. Hofmeister seriens effekt på opløseligheden blev bekræftet for både kationer og anioner, og for såvel B_{22} som for enzymopløseligheden ved pH 9 blev der observeret et minimum ved natriumnitrat og natriumthiocyanat koncentrationer mellem 0,1 M og 0,2 M. Sammenhængen mellem opløslighed og B_{22} er før blevet etableret af f. eks. Haas-Drenth-Wilson modellen, som i denne afhandling blev konstateret at være i kvalitativ overensstemmelse med de observerede tendenser for BHA.

Zusammenfassung der Dissertation

Ziel dieser Dissertation ist es, traditionelle sowie neue Ansätze zur Charakterisierung und Quantifizierung der Einflüsse verschiedener Lösungseigenschaften wie Salzkonzentration, pH und der Temperatur auf die Löslichkeit, das Kristallwachstum und das gesamte Phasendiagram des entsprechenden Proteins unter prozess-relevanten Bedingungen zu bewerten. Die getesteten und entwickelten Methoden wurden daraufhin optimiert, den Proteinbedarf zu reduzieren. Im experimentellen Teil dieser Dissertation wurden Konzentrate zweier rekombinanter α-Amylasen von *Bacillus halmapalus* (BHA) und *Bacillus licheniformis* (BLA) verwendet, die nur bis auf technischen Grad aufgereinigt wurden.

Die Entwicklung optimierter mikrobieller Stämme zusammen mit Fortschritten in Fermentierungstechnologien führen zu klareren industriellen Fermentationsmedien mit hohen Produktkonzentrationen, was die ökonomische Produktion von low-value-Proteinen begünstigt. Dabei steigt jedoch das Risiko, oberhalb der Löslichkeitsgrenze zu operieren, was zu unkontrolliertem Ausfällen oder Kristallisieren des produzierten Proteins führen kann. Daraus können Probleme während des Prozesses entstehen, wie z.B. der Verlust des Produktes in Unit-Operations wie Zentrifugieren oder Filtrieren, die zum Entfernen der Biomasse dienen. Der Betrieb von Aufreinigungsprozessen an der Löslichkeitsgrenze ist daher besonders kritisch, nichtsdestotrotz basiert deren Auslegung weitgehend auf trial-and-error, da umfassende Löslichkeitsdaten unter prozess-relevanten Bedingungen nicht verfügbar sind. Informationen über Löslichkeitseigenschaften von Proteinen sind daher nicht nur für die Auslegung und den Betrieb von Aufreinigungsprozessen von hohem Wert, sondern auch in frühen Screening-Routinen zur Ermittlung von potentiellen Molekülen mit gewünschten Eigenschaften. Löslichkeitsdaten sind nur für eine begrenzte Anzahl von Proteinen verfügbar, da traditionelle Methoden zur Bestimmung von Löslichkeitskurven viel Protein verbrauchen und die Kristallisierbarkeit innerhalb eines akzeptablen Zeitrahmens voraussetzen, was für viele Proteine nicht gegeben ist (Kapitel 1).

Die Löslichkeitseigenschaften von BHA werden im zweiten Kapitel intensiv charakterisiert. Die Löslichkeitskurven wurden mit Hilfe klassischer, temperaturkontrollierter Batch-Kristallisationen bestimmt, welche in Eppendorf-Reaktionsgefäßen bei einem Arbeitsvolumen von 1 mL durchgeführt wurden. Der Einfluss der Temperatur, des pH-Wertes und von

ausgewählten Anionen und Kationen der Hofmeister-Serie auf die Löslichkeit wurde quantifiziert. Es stellte sich heraus, dass die Löslichkeit nahezu unabhängig von der Temperatur war, jedoch sehr sensibel auf Änderungen des pH-Wertes reagierte. Die Hofmeister-Serie für Anionen konnte bestätigt werden, was mit Ausnahme von Lithium ebenfalls für monovalente Kationen zutraf. Es war erwartet worden, dass Lithium das schlechteste Ausfällungsmittel wäre, allerdings war es das beste. Die Bestimmung des Zeta-Potentials konnte zeigen, dass Lithium den isoelektrischen Punkt (pI) von BHA in Richtung höherer pH-Werte verschieben konnte, was das unerwartete Löslichkeitsverhalten des Enzyms in Gegenwart dieses Kations erklären könnte.

Im dritten Kapitel werden die Batch-Kristallisationsprozesse von BHA ausgehend von drei unterschiedlichen Übersättigungsgraden genauer untersucht. Die Kristallgröße, die Kristallgrößenverteilung, die Kristallkonzentration sowie die Proteinkonzentration im Überstand wurden als Funktion der Zeit bestimmt. Eine nicht zu vernachlässigende Anzahl von Kristallen bildete sich bereits während des Einstellens des pH-Wertes zur Erzeugung der Übersättigung. Während der ersten zwei bis drei Stunden des Prozesses stieg die Kristallkonzentration, blieb dann aber für eine gewisse Zeitdauer konstant, deren Länge im Wesentlichen von der anfänglichen Übersättigung bestimmt wurde. Nur wenn die Übersättigung hoch genug war, stieg die Kristallkonzentration weiter an und erreichte ein Gleichgewicht während der experimentellen Laufzeit von maximal 48 Stunden. Im Gegensatz zur Kristallkonzentration änderte sich die Größenverteilung der Kristalle während des Prozesses nicht, entsprechend blieb der mittlere Kristalldurchmesser nahezu konstant und bewegte sich zwischen 4 und 5 μm, so dass kein kontinuierliches Kristallwachstum beobachtet werden konnte. Die gemessenen Größenverteilungen der Kristalle waren im Wesentlichen unabhängig von der anfänglichen Übersättigung.

Der Einfluss ausgewählter Kationen und Anionen von der Hofmeister-Serie auf die Löslichkeit von BLA wird im vierten Kapitel behandelt. Im Gegensatz zu BHA ist BLA auf beiden Seiten des isoelektrischen Punktes stabil, so dass die Hypothese der Umkehr der Hofmeister-Serie in Abhängigkeit der Polarität des Proteins getestet werden konnte. Die Löslichkeit wurde bei pH 6, 7 und 8 bei Salzkonzentrationen von 0,1 und 0,5 M gemessen. Das Vorzeichen der Proteinladung wurde mit Hilfe des Zeta-Potentials bei oben genannten pH-Werten und Salzkonzentrationen von 0 und 0,1 M bestimmt. Die Bestimmung des Zeta-Potentials wurde bei Salzkonzentrationen über 0,1 M ungenau, so dass der pI bei 0,5 M nicht

bestimmt werden konnte. Mit Ausnahme von 0,5 M Natriumthiocyanat wurde ein Löslichkeitsminimum bei pH 7 gemessen. Der Effekt der Anionen auf die Löslichkeit von BLA folgte der Hofmeister-Serie. Eine Umkehr der Hofmeister-Serie für Kationen und Anionen in Abhängigkeit des Vorzeichens der Proteinladung konnte nicht eindeutig gezeigt werden.

Das Potential von Mikrotiterplatten zur Generierung von BHA-Phasendiagrammen bestehend aus Präzipitats- und Nukleationszonen sowie metastabilen und ungesättigten Zonen wird in Kapitel 5 bewertet. Temperatur-kontrollierte Mikrobatch-Experimente wurden bei einem Arbeitsvolumen von 200 µL durchgeführt und der gebildete Niederschlag mit Hilfe von Licht-Mikroskopie analysiert. Die Bildung von amorphen Präzipitat und Kristallen verschiedener Beschaffenheit wurde den entsprechenden Anfangsbedingungen zugeordnet, wodurch mehrere, das Niederschlagsverhalten beschreibende Zonen definiert werden konnten. Die Mikrotiterplatten wurden nach vier Tagen analysiert und die Löslichkeitsdaten, die im zweiten Kapitel präsentiert werden, mit in die Bestimmung der Phasendiagramme einbezogen, die als Funktion des pH-Wertes, der Natriumchlorid-Konzentration bei pH 7 und pH 9 sowie der Natriumthiocyanat-Konzentration bei pH 7 bestimmt wurden. In allen Fällen wurde eine weite Nukleationszone ermittelt. Nur bei sehr hohen Übersättigungen kam es zur Bildung von amorphen Präzipitat. Mit Natriumchlorid bei pH 9 wurde überhaupt kein Präzipitat unter den getesteten Bedingungen gebildet. Nach Beendigung der Experimente war die metastabile Zone sehr schmal für Natriumchlorid bei pH 9 und als Funktion des pH-Wertes in Abwesenheit zugesetzter Salze. Im Gegensatz dazu erweiterte sich die metastabile Zone mit ansteigender Natriumchlorid-Konzentration und in einem größeren Umfang ebenfalls mit ansteigender Natriumthiocyanat-Konzentration, jeweils bei pH 7.

Auch wenn der Proteinbedarf durch die oben präsentierten Methoden deutlich vermindert werden konnte, stellen die Verfügbarkeit ausreichender Proteinmengen und deren Kristallisierbarkeit deutliche Hindernisse dar, die oft nicht überwunden werden können. Daher werden im sechsten Kapitel weitere Methoden untersucht, die diesen Problemen entgegenwirken können. Die Bestimmung des zweiten osmotischen Virialkoeffizienten (B_{22}) stellt eine attraktive Alternative zu Löslichkeitsmessungen dar, da die entsprechenden Experimente in ungesättigten Lösungen durchgeführt werden können, so dass nur wenig Protein benötigt wird und auf zeitintensive Kristallisationsprozesse verzichtet werden kann. Der B_{22} stellt einen Bolzmann-gewichteten mittleren Größenwert der Protein-

Wechselwirkungen dar, bei dem positive B_{22}-Werte abstoßende und negative B_{22}-Werte anziehende Wechselwirkungen anzeigen. Somit kann eine Voraussage über Phasentrennungs- und Aggregationsprozesse gemacht werden. Die Anwesenheit bedeutender Mengen von löslichen Aggregaten konnte durch dynamische Lichtstreuungsexperimente (DLS) an BHA-Lösungen nachgewiesen werden. Daraus wurde gefolgert, dass diese Aggregate der Grund war, weshalb statische Lichtstreuungsexperimente (SLS) nicht zu reproduzierbaren Bestimmungen des B_{22} von BHA-Lösungen, wie hier verwendet, führten. Self-interaction chromatography (SIC) stellt eine Alternative zu SLS zur Bestimmung des B_{22} dar, die zu quantitativ gleichen B_{22}-Werten führt wie SLS, jedoch mindestens eine Größenordnung effektiver in Bezug auf Protein- und Zeitbedarf ist. Die in Kapitel 6 präsentierten Experimente bestätigten die Unempfindlichkeit von SIC Aggregaten gegenüber. Eine große Datenmenge von B_{22}-Werten wurde generiert und eine hohe Übereinstimmung zwischen B_{22} und Löslichkeitsmessungen für BHA festgestellt. Die Hofmeister-Serie wurde für Kationen und Anionen bestätigt und ein Minimum für B_{22} als auch für die Löslichkeit von BHA für Natriumnitrat und Natriumthiocyanat zwischen 0,1 und 0,2 M bestimmt. Eine empirische Korrelation zwischen der Löslichkeit und des B_{22}, wie z.B. durch das Haas-Drenth-Wilson-Modell, offenbarte eine qualitative Übereinstimmung mit den experimentell gefundenen Trends zwischen den beiden Größen.

Thesis summary

The overall aim of this thesis was to evaluate traditional and novel approaches for the characterisation and quantification of the influence of different solution properties such as changes in salt concentration, pH and temperature on protein solubility, crystal growth and the overall protein phase diagram under process relevant conditions. The tools were evaluated with the aim of reducing protein demand whilst operating in the presence of impurities. Concentrates of two recombinant α-amylases of *Bacillus halmapalus* (BHA) and *Bacillus licheniformis* (BLA) which had been purified only to technical grade were used in the experimental part of this thesis.

The development of optimised microbial strains along with advanced fermentation technologies leads to cleaner industrial fermentation broths with high product concentrations, facilitating the economical production of low-value-proteins in large-scale. At the same time, the risk of operating above the solubility limit and consequent uncontrolled formation of amorphous precipitation or crystals is significantly enhanced and can lead to substantial problems during processing, e.g., the loss of product during commonly employed biomass removal steps such as centrifugation or filtration. The design and the operation of downstream processes under conditions close to the solubility limit are particularly challenging and often rely on trial-and-error due to the absence of comprehensive solubility data for process relevant conditions. Thus information on the solution properties of proteins are of far reaching value not only for the design and operation of recovery processes but also in initial screening routines for promising new candidate molecules. However, solubility data are available for a limited number of proteins only, since the traditional determination of solubility curves requires a substantial amount of protein and that their crystallisation is possible within a reasonable time frame, constraints which are often not met (Chapter 1).

In Chapter 2 solubility properties of BHA were extensively studied. Solubility curves were obtained by classical temperature-controlled batch crystallisation processes conducted in Eppendorf tubes at a working volume of 1 mL. The influence of temperature, pH and selected cations and anions from the Hofmeister series on the solubility was quantified. The solubility was found to be almost insensitive to temperature but strongly dependent on pH. The Hofmeister series for anions was followed in the correct order which was also true for

monovalent cations, with the exception of lithium which was expected to be the worst precipitant but found to be the best. Measurements of the zeta potential were conducted and have demonstrated that lithium increased the isoelectric point (pI) of BHA which could explain the unexpected solubility behaviour of the enzyme in the presence of this ion.

In Chapter 3 batch crystallisation processes of BHA started from three different supersaturations were further studied. The crystal size, size distributions and the crystal concentration as well as the protein concentration in the supernatant were measured as a function of time. A significant number of crystals already formed during pH-adjustment to induce the supersaturation. Within the first two to three hours of the process, the crystal concentration further increased but then remained constant for a certain period, the length of which was determined by the supersaturation. Only if the supersaturation was high enough, the crystal concentration further increased and reached equilibrium within the experimental run time of maximum 48 hours. Contrary to the crystal concentration, the crystal size distribution did not change during the process and the mean diameter of the crystals remained constant and ranged between 4 and 5 μm such that no continuous growth could be observed. The crystal size distribution was essentially independent of the supersaturation.

The influence of selected cations and anions from the Hofmeister series on the solubility of BLA is presented in Chapter 4. In contrast to BHA, BLA is stable on both sides of the pI, allowing testing of the hypothesis that the Hofmeister series is reversed depending on the sign of the protein net charge. The solubility of BLA was measured at pH 6, 7 and 8 at salt concentrations of 0 M, 0.1 M and 0.5 M. The sign of the protein net charge in the salt-solutions was determined by measuring the zeta potential, at the above mentioned pH-values in the presence of 0 M and 0.1 M salt only. Zeta potential measurements became inaccurate at salt concentrations higher than 0.1 M, thus the resulting pI at 0.5 M could not be determined. With the exception of 0.5 M sodium thiocyanate, a minimum in solubility at pH 7 was found for all of the measured conditions. The effect of anions on α-amylase solubility was observed to follow the Hofmeister series. A reversal of the Hofmeister series for cations and anions depending on the sign of the protein net charge could not conclusively be demonstrated.

The potential of microtitre plates to generate BHA-phase diagrams consisting of precipitation and nucleation zones, as well as metastable and undersaturated zones is demonstrated in Chapter 5. Temperature-controlled micro-batch experiments were conducted at working

volumes of 200 µL and the precipitation behaviour analysed by light microscopy. The formation of amorphous precipitate and crystals of different habit was related to the corresponding starting conditions allowing different zones describing the precipitation behaviour to be defined. Inspection of the microtitre plates was conducted after four days and the solubility data presented in chapter 2 were incorporated. Phase diagrams were recorded as a function of pH, and of the concentration of sodium chloride at pH 7 and 9 and of sodium thiocyanate at pH 7, respectively. In all cases, a wide nucleation zone could be identified. Amorphous precipitate was formed only at very high initial supersaturations. For sodium chloride at pH 9, no amorphous precipitate was found. At the termination of the experiment, the metastable zone was found to be very narrow for sodium chloride at pH 9 and as a function of pH in the absence of added salts. In contrary, the metastable zone broadened with sodium chloride concentration and even more with sodium thiocyanate concentration, both at pH 7.

Although the protein demand could be significantly be minimised by the methods presented above, the availability of sufficient protein and the crystallisability may still remain constraints which are hard to overcome, thus other methods were examined in Chapter 6 using BHA as model protein. Determination of the second osmotic virial coefficient (B_{22}) is an attractive alternative to solubility measurements and offers the distinct advantage that experiments can be conducted in undersaturated solutions so that little protein and no time-consuming crystallisation processes are needed. The B_{22} represents a Boltzmann-weighted average measure of the protein-protein interactions where positive B_{22}-values correspond to repulsive interactions, while negative values correspond to attractive interactions, thus allowing for a prediction of phase separation and aggregation processes. Dynamic light scattering (DLS) of the BHA preparation used during the experimental work of this thesis demonstrated the presence of significant amounts of soluble enzyme aggregates which was concluded to be why static light scattering (SLS) did not lead to a reliable determination of B_{22} values of BHA solutions as used in this study. An alternative method to SLS is self-interaction chromatography (SIC) which has previously been shown to provide B_{22}-values which agree quantitatively with those made by SLS but is at least one order of magnitude more efficient in terms of protein consumption and time needed. The experiments presented in Chapter 6 confirm the insensitivity of SIC towards aggregates. Furthermore, a large set of B_{22}-data was generated and good agreement between B_{22} and solubility measurements was found for BHA. The Hofmeister series was confirmed for both cations and anions and a

minimum in B_{22} as well as in solubility at pH 9 was found for sodium nitrate and sodium thiocyanate between 0.1 M and 0.2 M. Correlations between solubility and B_{22} have previously been established e.g., by the Haas-Drenth-Wilson model which in this present thesis was found to be in qualitative agreement with the observed trends for BHA.

List of abbreviations

BHA	*Bacillus halmapalus* α-amylase
BLA	*Bacillus licheniformis* α-amylase
DLS	Dynamic light scattering
EDC	1-ethyl-3-(dimethylamino)propyl carbodiimide
ESEM	Environmental scanning electron microscopy
ESZM	Electrical sensing zone method
HDW	Haas-Drenth-Wilson model
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
IEF	Isoelectric focusing
LALLS	Low angle laser light scattering
MES	2-(N-morpholino)ethanesulphonic acid
MHB	MES-HEPES-Boric acid buffer
MTP	Microtitre plate
NHS	N-hydroxysuccine imidine
PDB	Protein Database
pI	Isoelectric point
PNP	p-nitrophenol
SDS PAGE	Sodium dodecyl sulphate - Polyacryl amide gel electrophoresis
SIC	Self-interaction chromatography
SLS	Static light scattering

Nomenclature

a_i	Activity coefficient of specie i	(-)
A_S	Total available surface	(m^2)
B_{22}	Second virial coefficient	(mol mL g^{-2})
B_{22}^{HS}	Hard sphere contribution	(mol mL g^{-2})
C	Mass protein concentration	(kg (kg $_{water}$)$^{-1}$)
C'	Molar protein concentration	(mol m^{-3})
C_{eq}	Mass protein concentration at saturation/equilibrium	(kg (kg $_{water}$)$^{-1}$)
C_P	Protein concentration	(mg mL^{-1})
D	Diffusion coefficient of protein molecule	(m^2 sec^{-1})
d (1,0)	Mean number diameter	(m)
d (4,3)	Mean volume diameter	(m)
d(H)	Hydrodynamic radius	(m)
d_0	Molecule diameter	(m)
d_c	Critical diameter of nucleus	(m)
f(κa)	Henry factor	(-)
f*	Monomer attachment frequency	(sec^{-1})
ΔG (V)	Free energy difference between a protein in solution and incorporated in the solid phase	(J)
ΔG (r)	Gibbs free energy	(J)
G	Overall growth rate	(m sec^{-1})
G_D	Growth rate controlled by transport	(m sec^{-1})
G_{SD}	Screw dislocation growth rate	(m sec^{-1})
J	Nucleation rate	(m^{-3} sec^{-1})
k	Boltzmann constant	(J K^{-1})
K	Constant for the calculation of the surface energy	(-)
k'	Chromatographic retention factor	(-)
K_1	Equilibrium constant of a protein monomer upon monomer addition	(-)
K_∞	Equilibrium constant of a protein aggregate upon monomer addition	(-)
k_d	Kinetic parameter of growth rate controlled by transport	(m sec^{-1})
$k_{dissolution}$	Kinetic parameter of dissolution	(J (molecule K)$^{-1}$)
K_{SD1}	Burton Cabrera Frank growth rate parameter I; SD: screw dislocation	(m h^{-1})
K_{SD2}	Burton Cabrera Frank growth rate parameter II; SD: screw dislocation	(-)
M_W	Molar weight of protein	(g mol^{-1})
n*	Work necessary for the formation of a nucleus	(J)
n_0	Refractive index	(-)
N_A	Avogadro number	$(mol)^{-1}$
P	Constant for the calculation of the surface energy	(-)
R	Gas constant	J $(mol\ K)^{-1}$
R_θ	Rayleigh ratio	(-)

r	Radius of a spherical nucleus	(m)
r*	Radius of the critical nucleus	(m)
S	Supersaturation ratio	(-)
T	Absolute temperature	(K)
t	Time	(sec)
U_E	Electrophoretic mobility	$(m^2 (V\ sec)^{-1})$
V_0	Retention volume no interaction	(mL)
V_m	Molar volume of protein	$(m^3\ mol^{-1})$
V_r	Retention volume upon interaction	(mL)
W	Potential of mean force	(J)
W*	Work necessary for the formation of a nucleus	(J)
z	Zeta potential	(mV)
z	Zeldovitch factor	(-)

Greek letters

α	volume shape factor	(-)
β	volume shape factor	(-)
ε	dielectric constant	(F m^{-1})
ϕ	phase ratio	(m^2/m^3)
γ	surface energy crystal/solution	(J m^{-2})
γ	interfacial free energy	(J)
η	dynamic viscosity	(Pa sec)
λ	wave length	(m)
μ_0	moment of order 0	(#)
μ_1	moment of order 1	(m)
μ_2	moment of order 2	(m^2)
μ_3	moment of order 3	(m^3)
μ_4	moment of order 4	(m^4)
μ_i	chemical potential of specie i	(m^4)
μ_k	moment of order k	(-)
ν_0	volume occupied by a molecule	(m^3)
θ	time without dimension	(-)
ρ	protein density	(kg m^{-3})
ρ_c	density of solid	(kg m^{-3})
ρ_S	number of immobilised protein per unit area	(molecule m^{-2})
Ω	molecular volume occupied by one growth unit	(-)
Ω_1, Ω_2	normalised vectors describing the angular position and orientation of the protein molecules	(-)

Table of Contents

1 Introduction

1.1 Downstream processing of industrial enzymes

1.1.1 General aspects

Micro organisms are becoming more and more attractive for the production of large quantities not only of enzymes but also of small organic molecules as well as of protein biopharmaceuticals. The vast majority of industrially produced enzymes are of bacterial or fungal origin. The maximisation of the fermentation yields and the formulation of high strength products constitute core issues of the economical production of industrial enzymes. This development is boosted by rapid advances in recombinant DNA technologies which enables the stable and safe production of enzymes by inserting the gene encoding the desired protein into the host micro organism. This host micro organism should be easy to ferment, non-pathogenic, easy to be genetically modified, and should produce the protein of interest in high yields at minimised costs. It is generally beneficial for the downstream process if the enzyme of interest is produced in an extracelluar manner such that costly cell disruption procedures, e.g., by homogenising, bead mills or by enzymatic or chemical cell lysis can be circumvented. After termination of the upstream process (Figure 1.1), the fermentation broth typically consists of between 80 and 99% of water. The product of interest is very often present in small amounts between 1 and 20% together with a complex mixture of by-products and residual additives and substrates (Hanko & Rohrer, 2004). Apart from separating the product from the cell and other solids, additional purification steps are typically required, the complexity of which depends on the product and application. Smaller product titres generally require a larger number of unit operations in the downstream processing which leads to the more costly products. Directed metabolic and genetic engineering has increased the production level into the range of grams per litre; strains of *Aspergillus* have been reported to produce titres of more than 30 mg/mL protein of high homology (Carlsen & Nielsen, 2001) which may reduce the complexity of the subsequent downstream process. However, the treatment of the fermentation broth is complicated by the risk of microbial contamination and enzymatic, thermal, chemical and stress-related degradation and by considerable batch-to-

batch variations (Jacobsen et al., 1998). As a consequence, the choice of temperature, pH, additives and mixing is thus restricted by constrains dictated by the limited stability of the product.

Whether or not cell disruption is necessary (Figure 1.1), the cells and/or debris are subsequently removed by either centrifugation or filtration and the product is concentrated by ultrafiltration, extraction, precipitation or, in case of thermostable products, evaporation. The degree of purity so obtained may already be sufficient for low value products such as industrial enzymes but in case of pharmaceutical products, more specific steps like size exclusion, hydrophobic interaction, affinity or ion-exchange chromatography, the latter probably being the most commonly used technique in industrial purification (Staby et al., 1998), are mandatory to ensure the required product purity and quality (Figure 1.1). The number and sequence of the unit operations have to be adjusted to every product and application. In general, high capacity-low cost operations (e.g., centrifugation) are performed in early stages and low capacity-high cost/high resolution steps are placed in the later stages of the downstream process. Product formulation, labelling, packing, storage and transportation are steps following the downstream process which may also pose important constraints to be considered in the product and process development.

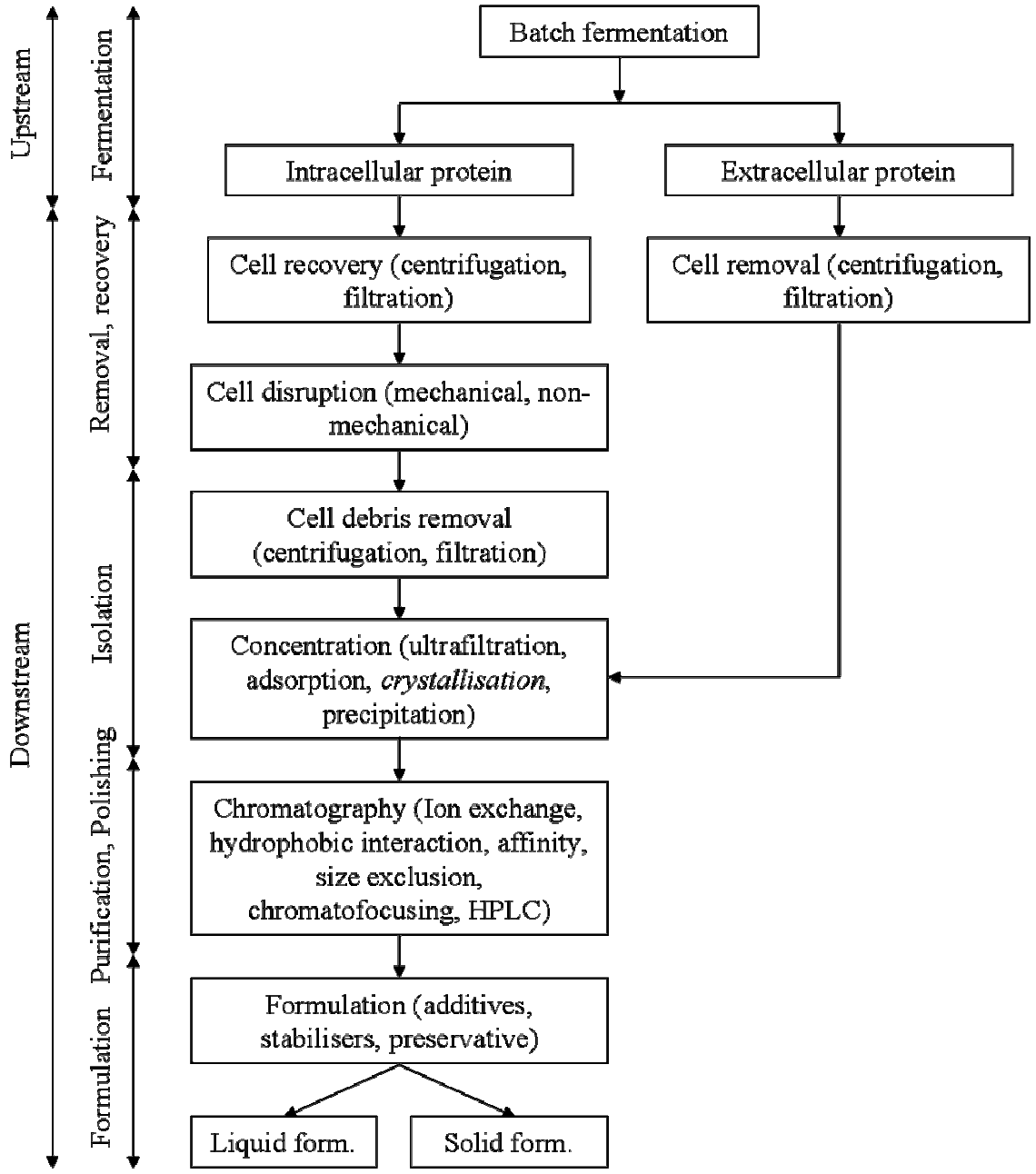

Figure 1.1 Generalised process flow sheet for the manufacture of soluble intra- or extracellular proteins (Walsh & Headon, 1994).

Regrettably, yield and purity are often opposing each other. The higher the purity demand, the more unit operations are typically necessary which reduces the product yield (Schügerl, 2000). An imaginary process consisting of a total of 10 unit operations of a constant step yield of 90% would result in a final overall recovery yield of less than 40% (Fish & Lilly, 1984). Particularly the production of protein biopharmaceuticals generally requires significantly more than 10 unit operations; as an example, the recombinant production of insulin in *E. coli*

has been reported to require 27 unit operations with the consequences mentioned above (Prouty, 1991). Another example is the purification of a *Penicillium citrinum* lipase in which a factor of 379 was achieved by five unit operations resulting in a highly homogeneous product at an overall yield of only 15% (Krieger et al., 1999). Depending on the purity demand, downstream operations account for approximately 50 to 80% of the total manufacturing costs such that the reduction of the number of unit operations in the recovery process is probably the most effective way to reduce the final product price (Spalding, 1991).

1.2 Crystallisation in downstream processing

In light of the above, crystallisation is an interesting unit operation since it efficiently combines concentration and purification. Until recently, it was believed that high product purity was an essential requirement for successful crystallisation. As a consequence, crystallisation was assumed to be applicable only at very late stages of the downstream process and thus only occasionally employed. Recent studies have, however, demonstrated that high product purity prior to crystallisation may not always be necessary (Jacobsen et al., 1998; Judge et al., 1995). This would offer the distinct advantage of synchronised and effective purification and concentration applicable at almost any stages of the downstream process and could lead to very efficient and cheap product capture.

As the crystallisation process is very specific, unwanted side activities, e.g.,, proteolysis, and colour can concurrently be removed. The advent of less complex media compositions as well as of highly productive strains which secret the product directly into the fermentation broth, enhances the feasibility of crystallisation as an integrated unit operation in protein downstream processing. In case of extracellular protein production, a downstream process only consisting of simple cell removal, filtration, concentration and crystallisation could probably lead to sufficient product purity, e.g., for industrial enzymes; such a process would arguably be a very cheap and fast way of protein manufacture difficult to surpass.

The growing demand for products of high strength and the need for efficient use of the product facilities enforce operations near the solubility limit which enhances the probability of uncontrolled crystallisation or precipitation during fermentation or subsequent downstream operations, particularly when the product is of low solubility (Buque-Taboada et al., 2004).

Due to the absence of comprehensive solubility data and phase diagrams at process relevant conditions, the design and the operation of downstream processes often rely on trial-and-error. Information on the solution properties would thus be of high value but is available for a limited number of proteins only, since the traditional determination of solubility curves requires a substantial amount of protein which can be crystallised within a reasonable time frame, constraints which are often not met. The true value of phase diagrams is not only restricted to finding favourable crystallisation conditions. Since protein crystallisation is almost exclusively conducted in batch mode, it may not be the ideal unit operation, particularly in high throughput processes. In such a case, phase diagrams provide information how to run processes avoiding any kind of undesired phase transitions. Only the availability of phase diagrams would pave the way to a complete process-in-control large scale enzyme production.

Protein crystallisation is an old technique, first employed by Hünefeld in 1840 for the crystallisation of haemoglobin. It has mainly been used for the determination of tertiary structure of proteins aiming for high quality crystals to obtain sufficient resolution (Klyushnichenko, 2003). In spite of the dramatic increase in successfully crystallised proteins judged by the number of resolved structures inserted in the Protein Data Base (PDB), insulin is one of very few protein biopharmaceutical sold in crystalline form which has gained marketing authorisation (Brange et al., 1988; Prouty, 1991). The number of approved biopharmaceuticals is rapidly increasing but problems such as limited stability and administering options remain. Crystalline proteins are advantageous as they often possess a high bioavailability, are easier to handle and are very stable, since physical and chemical degradation is significantly reduced, thereby maintaining the biological integrity over a long time. Moreover, they are beneficial in the delivery of biopharmaceuticals to achieve high concentration, low viscosity formulation and controlled release delivery (Basu et al., 2004). The contradiction between increasingly successful crystallisation trials for structure resolution and the very low number of crystalline proteins produced in large scale may be explained as follows: The vapour diffusion methods typically employed such as hanging or sitting drop are not relevant for large-scale crystallisation since they aim for large single crystals of high quality and are not scalable. The choice of chemicals to be used in large scale crystallisation is limited, not only for economical reasons but also because of possible difficulties in the downstream process. Next to the often encountered resistance of proteins to crystallise in larger scale, industrial protein crystallisation may face additional challenges: Due to the large

voids and channel, the solvent content of protein crystals can range between 20 and 90% (Chernov, 1997). As a consequence, the crystals are more fragile than inorganic crystals and are therefore more difficult to handle. Moreover, protein crystals typically contain large amounts of dissolved impurities of low molecular weight; thus the efficiency of crystallisation processes to concentrate and purify proteins is reduced compared to inorganic molecules (but generally still higher than other concentration steps typically employed in protein downstream processing). In addition, the degree of supersaturation necessary to induce nucleation and subsequent crystal growth is higher for protein crystals (1.2 to 100) compared to 1.1 used in industrial crystallisation of inorganic molecules (Durbin & Feher, 1986; Feher, 1986) which may require costly concentration steps prior to crystallisation. Nucleation and growth rates of protein crystals are significantly reduced which leads to longer processing times with negative effects on the economy of the production (Rosenberger et al., 1996). Finally, protein crystals are smaller than crystals of inorganic molecules which is per se not a disadvantage as long as they are homogeneous in size. However, bigger crystals are often easier to harvest since they have e.g., a better filterability due to reduced pressure loss (Rohani et al., 1990).

1.2.1 Phase diagrams

For the design of efficient downstream processes involving crystallisation steps, a fundamental understanding of the kinetics and the phase behaviour is required. Phase diagrams are subdivided into the precipitation zone where amorphous precipitate is formed, the nucleation zone which leads to crystal formation and growth, the metastable zone where no crystals are formed but existing crystals will grow and the undersaturated zone where no crystallisation can occur (Figure 1.2). To understand the properties of the phase diagram, it is necessary to consider that the crystallisation process consists of nucleation and growth which will be further discussed in the following.

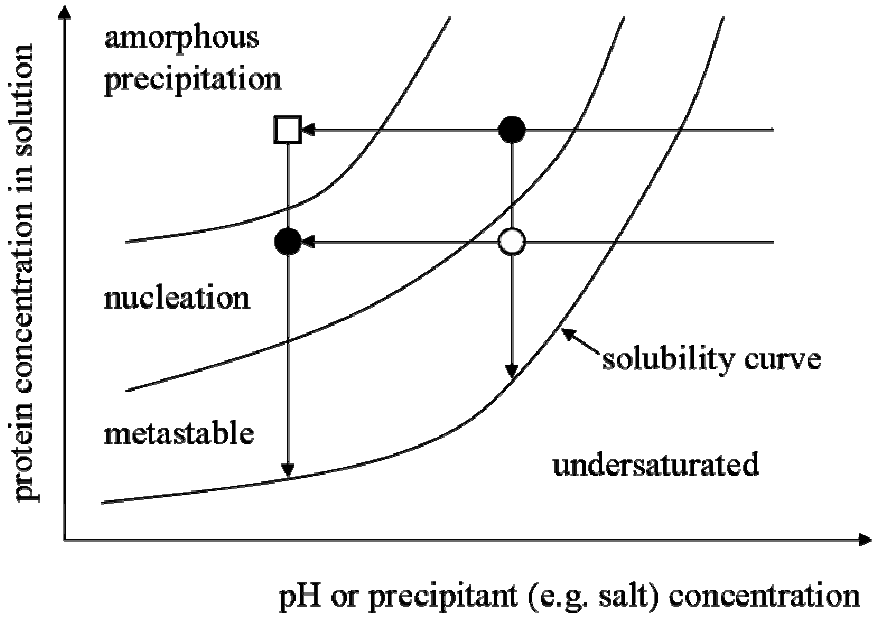

Figure 1.2 Schematic phase diagram of a given protein. For the determination of the solubility curve it is mandatory that crystals are formed since they represent the only solid phase in true thermodynamic equilibrium with the liquid phase. Accordingly, batch crystallisation experiments have to be conducted such that the system is moved across the solubility line into the nucleation zone (filled circles). In case the system is moved into the precipitation (open square) or metastable zone (open circle), the experiments have to be repeated. Unordered amorphous precipitate cannot be considered as a true solid phase, thus it cannot be in a thermodynamic equilibrium with the liquid phase. In the metastable zone, the supersaturation is so low that no crystals will appear within a reasonable time.

1.2.2 Nucleation

Nucleation of a new solid phase originates from random energy fluctuations around the increased mean value of the free energy induced by the supersaturation. Nucleation is thus a stochastic event. Several types of nucleation can be defined: When the nucleation takes place in an environment completely free of crystals, it is termed primary nucleation and is homogeneous when it occurs spontaneously in the bulk of the solution independent of the presence of solid phases like dust or surfaces. Primary nucleation induced by foreign particles such as dust or other impurities is referred to as heterogeneous. This type of nucleation is practically unavoidable since impurities to some degree will always be present. Secondary nucleation implies the nucleation in the presence of already formed crystals. The most commonly proposed reasons for secondary nucleation are crystal fracture and attrition. Fracture is most likely to affect crystals in strongly agitated systems where impacts between

two crystals or crystal and reactor wall as well as by crystal break-up induced by fluid stresses are very likely to occur. The size of the resulting fragments is in the same order of magnitude than the parental crystals. Attrition occurs mainly in systems of high crystal concentration due to crystal-crystal or crystal-reactor contacts. The fragments so obtained are much smaller than the parental crystals (Synowiec et al., 1993).

Crystals will not grow out of all supersaturated solutions. To create a new phase, the system must overcome a certain energy barrier. If the supersaturation is too low, the amplitude of the energy fluctuations is not high enough to exceed this energy barrier. The change in Gibbs free energy $\Delta G(r)$ upon formation of a spherical nucleus of radius r is composed of two terms, namely a surface and a volume term. The surface term refers to the energy which has to be provided to create a crystal surface unit whereas the volume, term the magnitude of which increases with supersaturation, describes the energy gain resulting from the decrease of the free energy of the system as

$$\Delta G(r) = 4\pi r^2 \gamma + \frac{4}{3}\pi r^3 \Delta G_V \qquad (1)$$

where γ is the interfacial free energy between a crystal nucleus and the bulk solution, r the radius of the nucleus and $\Delta G_V = \frac{kT}{\Omega}\ln S$ is the free energy difference between a protein molecule in solution and incorporated into the solid phase with Ω being the molar volume occupied by a growth unit (molecule) in the cluster and S the solution supersaturation, k the Boltzmann constant and T the absolute temperature (Gibbs-Thompson-expression). ΔG_V is negative for all supersaturated solutions whereas the surface term is always positive. In an undersaturated solution, both terms are positive such that crystallisation is thermodynamically not possible. In the metastable zone, the volume term is negative but the degree is not enough to exceed the contribution from the positive surface term. Consequently, no nucleation can proceed but existing crystals, for which the volume term is of higher magnitude than the surface term, can continue growing. On the borderline between metastable and nucleation zone the two terms counterbalance each other. In the nucleation and precipitation zone, the Gibbs free energy is always negative and the phase separation process thermodynamically favourable (Arakawa & Timasheff, 1985).

The radius at which equation (1) exhibits a maximum, is the radius of the critical nucleus, i.e. at this radius, surface and volume term are equal. Differentiation of equation (1) provides the radius of the critical nucleus r* as

$$r^* = \frac{2\Omega\gamma}{kT \ln S} \tag{2}$$

The size of the critical nucleus decreases with increasing supersaturation. At the same time the induction time to form the critical nucleus decreases with increasing supersaturation (McPherson et al., 1995) which is understandable as at higher supersaturations more protein molecules are available such that the probability for a molecule to get correctly introduced into the array forming the nucleus is increased.

The thermodynamic difference between crystallisation and precipitation can be expressed by the ratio of the equilibrium constants upon monomer addition to an existing aggregate K_∞/K_1 where K_∞ is the equilibrium constant of the aggregate and K_1 the equilibrium constant of the monomer. This ratio is very different depending on whether the protein precipitates in amorphous or crystalline form (Feher & Kam, 1985). For a compact crystalline configuration, the addition of a monomer creates more than two new bonds compared to a single bond for the creation of dimer from a monomer (García-Ruiz, 2003) and it follows that $K_\infty/K_1 >> 1$. These large ratios (for lysozyme crystals approximately 35) arise from the three-dimensional character of the crystals. In contrast, amorphous precipitate is approximated by one-dimensional chains of protein molecules to which monomers can be added only at the ends, so that only one bond is involved upon incorporation of a new molecule. Thus the incremental energy difference on adding a monomer is independent of the size of the aggregate and the competition between the volume and the surface energy term (equation 1) is eliminated and consequently, no energy barrier exists for linear growth, accordingly, $K_\infty/K_1 \sim 1$. In the absence of an energy barrier, the growth proceeds without time lag (Kam et al., 1978).

1.2.3 Crystal growth

The formation of the critical nucleus induces the phase of crystal growth. Two main mechanisms are underlying crystal growth, (i) mass transfer from the bulk of the solution to the crystal interface and (ii) attachment of the molecules into the crystal lattice (Chernov, 2003). Mass transfer is governed by diffusion and convection due to density driven gradients or by stirring. The attachment of molecules to the crystal lattice can be described by several mechanisms where the two seemingly most relevant for protein crystal growth are described in the following.

Two dimensional nucleation growth

The first mechanism is two dimensional nucleation growth, which is the simplest form of crystal growth. Here, molecules passing through the close proximity of the surface of a growing crystal get weakly adsorbed. They may join together to form small, two-dimensional islands and spread outward in a layer one molecule thick, with other islands forming and growing on top. In this dynamic growth process, molecules continually adsorb and dissolve from islands. Once the adsorbed molecules have formed a two-dimensional cluster on the surface which exceeds the critical size, the crystal growth continues. This cluster is then able to incorporate molecules colliding on the surface. As the step edges advance, single molecules may diffuse from islands in the vicinity of the outer edge of a step or from solution and be captured by that edge. In this way the edge acts as a sink to diffusing molecules. It should be remembered, however, that molecules have complex shapes that prevent them from bonding in every orientation. For example, a molecule may have to diffuse to the edge of a step many times until it has the correct orientation for incorporation. This mechanism is likely to be dominating at high supersaturations (McPherson et al., 1995).

Screw dislocation

The second mechanism is screw dislocation which causes a kink in the otherwise perfectly layered crystal surface into which the molecules can adsorb and thus be incorporated into the lattice. The growth then propagates in a spiral manner around the centre of dislocation. Screw dislocations may arise from the incorporation of an impurity or misaligned macromolecular building units and are more common at lower supersaturations (Malkin et al., 1996).

1.2.4 Growth cessation

In spite of its high relevance, growth cessation is probably one of the least understood aspects of protein crystallisation. Even in the presence of excess protein, many protein crystals do not exceed a certain size. Two different mechanisms can be discriminated. One mechanism is the gradual poisoning of the crystal surfaces by impurities which adsorb to the crystal lattice, thus preventing further growth (Weber, 1991). Typically, the impurities bind more weakly to the lattice than the native molecules and are thus excluded from the crystal surface. However, when the concentration of the native molecules falls below a certain value and the addition of impurities to the lattice becomes dominant, the growth stops. As an example, the crystallisation of haemoglobin C can be inhibited by haemoglobin F. The degree of this inhibition can be steered by the concentration of haemoglobin F (Hirsch et al., 1988). Another mechanism that can lead to growth cessation is the accumulation of building defects in the crystal lattice. Lysozyme crystals broken into smaller pieces grew to their original size (Kam et al., 1978), suggesting structural defects since they were propagated into the new crystals. Structural defects may be expected to be more common in rapidly growing crystals. The final size of lysozyme crystals have been found to be highly dependent upon the rate at which the critical supersaturation was approached (Gernert et al., 1988). This finding may be of high relevance in large scale batch crystallisation processes. If e.g., the supersaturation is induced by changes in pH, the way either the acid or base is added may have a big impact on the resulting crystal size.

1.2.5 Protein batch crystallisation

Although inorganic molecules could be crystallised in batch and continuous mode, batch crystallisation is extensively used in the chemical industry, mainly because of its simplicity, flexibility and the manageable development work and capital investment (Toyokura, 1995). These advantages are also believed to be true for the large-scale protein production. Moreover, the slow nucleation and growth kinetics limit the feasibility of continuous protein crystallisations. In a batch crystallisation process, the component to be crystallised first has to be solubilised. Subsequently, the solution conditions have to be changed such that the system is moved from undersaturation across the solubility curve to become supersaturated. Supersaturation is typically induced by changing parameters such as pH, ionic strength or

temperature. After a certain lag phase, which depends on the degree of supersaturation and is typically much longer for proteins compared to inorganic molecules, the crystallisation will start (Figure 1.2). This process can be subdivided into nucleation and crystal growth but happens simultaneously in a batch crystallisation. At the beginning of the crystallisation process, i.e. at high supersaturation, the nucleation is the dominating process; with decreasing supersaturation crystal growth is becoming more important until the supersaturation is degraded, the equilibrium is reached and the process is finished. In a classical batch process, the system remains closed and all parameters are kept constant after induction of the supersaturation. In order to obtain large crystals homogeneous in size, which is often desirable (Rohani et al., 1990), the nucleation rate should be low to create a limited amount of crystals, and the growth rate high to quickly finalise the process. By seeding, i.e. the introduction of already existing crystals into the metastable zone, nucleation and crystal growth can completely be decoupled (Bergfors, 2003). Other strategies to control size and size distribution of the resulting crystals are e.g., constant temperature control or constant supersaturation control (Shi et al., 2005). These methods may, however, be of limited practical relevance for large-scale protein crystallisation due to the generally slow response of the system to changes in process parameters and due to the absence of phase diagrams and tools e.g., to control the degree of supersaturation.

1.3 Factors influencing crystallisation

1.3.1 Protein

Probably the by far most important parameter is the protein to be crystallised. The Protein Data Base (PDB) provides a plethora of information under which conditions proteins have been crystallised. It is conspicuous that very few proteins could be crystallised under the same conditions which is an indication of the huge differences in the chemical and physical properties proteins exhibit. Proteins differ in number and sequence of amino acids which has an impact on their molecular weight, the three-dimensional structure, the biological activity, the isoelectric point and the distribution between hydrophobic and hydrophilic patches on the surface, just to name a few. The replacement of only one amino acid by another one could already cause dramatic changes in the properties mentioned above (Nielsen & Borchert,

2000). It is to date not possible to derive general guidelines leading to successful crystallisation, e.g., based on the amino acid sequence of a given protein (D'Arcy, 1993).

1.3.2 Protein concentration

Since crystallisation only takes place in supersaturated solutions, it is obvious that the protein concentration is an essential parameter. It is thus mandatory that the protein can be solubilised which may sound like a banal requirement but is far from self-evident for strongly hydrophobic membrane proteins which is often the main reason for their high resistance to crystallisation (Fritsch et al., 2002); they typically require the presence of non-ionic surfactants like Triton X to get solubilised (El Bawab et al., 1999). Moreover, it is necessary that the macromolecules maintain their native structure when dissolved to relatively high concentrations. On the other hand, too high protein concentrations result in non-specific aggregation which is obviously not desired either. Optimal concentration ranges leading to crystallisation are highly protein and condition specific and are subject to careful inspection. It should be considered that the protein concentration which determines the supersaturation ratio also influences the size of the resulting crystals which is commonly larger at lower supersaturations (Ataka & Tanaka, 1986). At the same time, lower supersaturation ratios can lead to significantly prolonged induction times which may often be unacceptable.

1.3.3 pH

The pH is probably the most important variable to be investigated in a search for crystallisation conditions (Giegé et al., 1995). At the pI the macromolecule carries an equal number of positive and negative charges and is the pH of lowest solubility. For a majority of proteins, the formation of amorphous precipitate seems to be particularly dominating over crystallisation at the pI. The size and morphology of resulting crystals often strongly depend on the pH. For lysozyme, tetragonal crystals are formed at pH 4.5 whereas orthorhombic crystals are obtained at pH 8 (Weiss et al., 2000). As already mentioned, larger crystals are obtained when grown at lower supersaturations (Ataka & Tanaka, 1986). Since the supersaturation can conveniently be controlled by the pH, its use to steer the crystal size is obvious.

1.3.4 Temperature

The influence of temperature on crystallisation can significantly differ from protein to protein. The solubility of apoferritin (Petsev et al., 2001) is thus independent on temperature, a direct dependence between temperature and solubility is seen with lysozyme (Pusey & Munson, 1991) whereas carbomonoxy haemoglobin C showed a retrograde behaviour (i.e. decreased solubility with increasing temperature) (Vekilov et al., 2002). Here it is also important to note, that irrespective of the direct influence on the solubility, the temperature is also important for the rate of crystal growth. An increased temperature provokes a higher Brownian motion leading to an enhanced contact either between protein monomers or between monomers and crystals which could result in faster nucleation and growth rates. It may have to be examined for each system in question whether or not changes in temperature lead to improvements e.g., on crystallisation times or crystal shapes. When the temperature is to be changed during a large scale crystallisation process, it should be considered that it may take a long time before the entire processed reactor volume has been adjusted to the new temperature. However, it should be born in mind that the temperature is a process parameter which within certain ranges can be arbitrarily reset without changing the solution composition.

1.3.5 Ion type and concentration

A wide variety of metal ions have been reported to promote or contribute to the crystallisation of different macromolecules. Often, the ions are essential for biological activity and may therefore be crucial in maintaining structural features of the molecule. Zink plays an important role in the insulin crystallisation (Kadima et al., 1993) and calcium is essential for the structural integrity of α-amylase (Nielsen et al., 2003). Next to the structural importance of some metal ions, anions are generally very effective in influencing the solubility and a characteristic order of efficiency by which they change the solubility, commonly referred to as the Hofmeister series can be established (Hofmeister, 1888). The Hofmeister series for anions is typically given as: sulphate > acetate > citrate > tartrate > chloride > nitrate > thiocyanate. The Hofmeister series for monovalent cations is stated as: lithium > sodium > potassium > ammonia > rubidium > caesium (Bénas et al., 2002). Ions on the lower end of this series are very efficient in lowering the solubility and are kosmotropes (structure makers), meaning that they are highly hydrated and commonly not very polarisable, sometimes also

called hard acids or bases (Pearson, 1963). Due to the shielding effect of the hydration layer around these molecules, they are excluded from the protein molecule such that a water-enriched zone in the proximity of the protein molecule is formed (preferential hydration) which is thermodynamically unfavourable. As a consequence, the association of proteins reduces the area of this zone, hereby their solubility is decreased. Ions on the higher end of the Hofmeister series are termed chaotropes (structure breakers) and are very polarisable (also referred to as soft acids or bases). They are not strongly hydrated and thus may have access to the protein surface and can change the solubility by binding to or screening of oppositely charged amino acid residues. Examples are nitrate and thiocyanate. Interestingly, the reversal of the Hofmeister series has been demonstrated for lysozyme with respect to their effect on solubility and was explained by the fact that given ions would bind differently to protein surfaces depending on the net charge of the protein (Riès-Kautt & Ducruix, 1989; Riès-Kautt & Ducruix, 1997). This reversal of the series has also been shown for human fibrinogen (Leavis & Rothstein, 1974) but it is still controversial whether it is a property generic to all proteins. Such a net charge dependent reversal of the series, if relevant for the protein of interest, would be of high relevance in processing in case salts are to be used to steer solubility.

1.3.6 Other factors

Other factors which will not be further discussed but may be influential for protein crystallisation can be the crystallisation mode (e.g., hanging or sitting drop, batch), pressure, gravity, surfaces, mixing, electrical or magnetic fields, viscosity, reductive or oxidative environment, detergent or surfactant concentration, purity of the macromolecule and the solution, posttranslational modifications such as phosphorylation or glycosylation, the source of the macromolecule, genetic or chemical modifications, stability of the macromolecule etc (Giegé et al., 1995).

1.4 The history and industrial relevance of enzymes

1.4.1 The history of industrial enzymes

Since the prehistoric time, mankind has been using fermentation processes e.g., for baking, brewing and alcohol production. There is evidence dating back to about 800 BC for the use of enzymes in the cheese production. The first scientific study on enzymes was reported in 1833 by Payen and Persoz who isolated an amylase complex from germinating barley[1]. Already two years later, it was discovered that starch can be broken down to glucose more efficiently by malt extracts than with sulphuric acid. The term enzyme was introduced by Kühne describing a substance located in yeast inducing fermentation (enzyme: Greek for "in yeast"). In 1894 the lock-and-key-model for the catalytic activity was proposed by Fischer which was based on the properties of glycolytic enzymes. In 1897 it was demonstrated that cell free extracts from yeast could break down glucose into ethanol and carbon dioxide. The fundamentals of enzyme kinetics date back to 1903 when Henri concluded that an enzyme combined with the matching substrate form an enzyme-substrate complex, hereby doing the groundwork for enzyme catalysis. Based on this idea, the general theory of enzyme action was described in mathematical terms by Michaelis and Menten in 1913. They postulated that the enzyme E first combines with its substrate S to form the enzyme-substrate complex ES in a relatively fast and reversible reaction. This complex breaks down into the product P and releases the enzyme again in a relatively slow but also reversible reaction. Only in 1926 it was discovered that enzymes are a type of protein. The use of enzymes in detergents which is to date their largest application started in the 1930s after Röhm filed a patent on the use of pancreatic enzymes in pre-soaked solutions[2]. Subtilisin, an alkaline bacterial protease, was the first important enzyme used in large scale for laundry detergents (van Ee, 1992).

[1] http://www.novozymes.com/cgi-bin/bvisapi.dll/discover/discover.jsp?cid=-9281&id=13226, February 2006

[2] http://www.novozymes.com/cgi-bin/bvisapi.dll/discover/discover.jsp?cid=-9281&id=13226, February 2006

1.4.2 The industrial relevance of enzymes

Many chemical processes suffer from severe disadvantages as the often non-specific reactions lead to poor product yields and unwanted by-products which may be difficult and expensive to dispose of. Chemical processes conducted under harsh and hazardous conditions such as high temperature, pressure, alkalinity or acidity may require expensive equipment and control systems. Unwanted by-products, low production yields, high chemical and energy consumption together with high equipment investments have a negative impact both on the economical and the ecological balance of a given process. In light of this, the use of enzymes may offer some benefits since enzyme reactions can be conducted under mild conditions, are highly specific, possess high reaction rates and are the product of fermentation processes of micro organisms using renewable resources, contributing to the sustainability of any given production process catalysed by enzymes (Nielsen & Borchert, 2000). Further advantages of enzymes are that they are very efficient so that only small amounts are required to initiate chemical reactions. Consequently, the use of solid and liquid enzyme formulations reduce the demand for storage space. Recent advances in genetic and protein engineering are leading to a constantly increasing number of industrial applications of enzymes since e.g., stability, economy and specificity could be substantially increased. Today enzymes are thus used in a number of industries including applications in the animal feed, baking, brewing, diary, detergent, fruit and vegetable processing, leather, fuel alcohol, personal care, pulp and paper, starch, sugar, textile and wine industry (Olsen & Falholt, 1998). In 1999, the market for technical enzymes (mainly for detergents and textile production) accounted for 63% followed by the food industry (mainly baking, beverage and dairy) with 31% and the feed industry of 6% of the total enzyme sales. In the same year, the world market for industrial enzymes was worth around € 1.5 billion, which in 2003 already had increased to € 1.8 billion[3, 4]. The market is expected to continue growing by 5% per year.

[3] http://freedonia.ecnext.com/coms2/summary_0285-293713_ITM, February 2006
[4] http://www.forbes.com/2001/11/07/1107gcor.html, February 2006

1.5 α-Amylases

α-Amylases (α-1,4-glucan-4-glucanohydrolase; EC 3.2.1.1.) are monomeric enzymes that catalyse the hydrolysis of the internal α-1,4-glycosidic bond in starch and related oligo- and polysaccharides (Henrissat, 1991). α-Amylases are so-called retaining glycoside hydrolases, which means that the reduced end of the reaction product will retain its α-configuration of the anomeric carbon during the hydrolysis.

1.5.1 Common features of α-amylase structures

The majority of α-amylases belong to group of enzymes related in sequence named glycoside hydrolase family 13 and consists of around 1700 members of which the structure of 40 has bee solved. Mammalian and bacterial α-amylases consist of three domains: A central $(\alpha/\beta)_8$ TIM-barrel (Triose phosphate isomerase; the enzyme for which this tertiary fold was first observed) forms the core of the molecule and contains the active site (domain A) (van der Maarel et al., 2002). An (α/β) barrel is formed by central twisted β-sheets surrounded by α-helices and is built up by β-α-β motifs, in which all the β-sheets are parallel. The subscript 8 indicates that the barrel consists of eight β-strands, which is the most commonly encountered number in these barrels. Domain B is a long complex loop protruding from the third β-strand and third α-helix of the barrel and the C-terminal and varies substantially in size and structure among the amylases (Nielsen & Borchert, 2000). Domain C contains a Greek motif (Suvd et al., 2001). This common motif is formed when β-strands align to form an antiparallel β-sheet. For most α-amylases, the C-domain is formed by eight stranded β-sheets. The shape is similar to a design found on Greek pottery, hence the name. Mammalian α-amylases are characterised by the presence of several disulphide bridges whereas they are generally not found in bacterial amylases (Machius et al., 1995). A conserved calcium binding site (Ca I) which is located at the interface between domains A and B is characteristic for all known α-amylases (Machius et al., 1995; Machius et al., 1998). The conserved calcium ion is very tightly bound to the molecule (Nielsen & Borchert, 2000) and plays a crucial role in structural integrity of the enzyme since it is too far away from the active site to participate directly in the catalysis (Vallee et al., 1959). Additional calcium binding sites have been identified for

Bacillus licheniformis α-amylase (Declerck et al., 1997) and *Bacillus halmapalus* α-amylase (Brzozowski et al., 2000). The conserved (Ca I) and the second calcium ion (Ca II) together with a sodium binding site form a linear Ca-Na-Ca arrangement, the so-called triad, characteristic of many bacterial α-amylases (Declerck et al., 2004). A third calcium binding site (Ca III) is situated at the interface between domains A and C where it acts as a bridging ion.

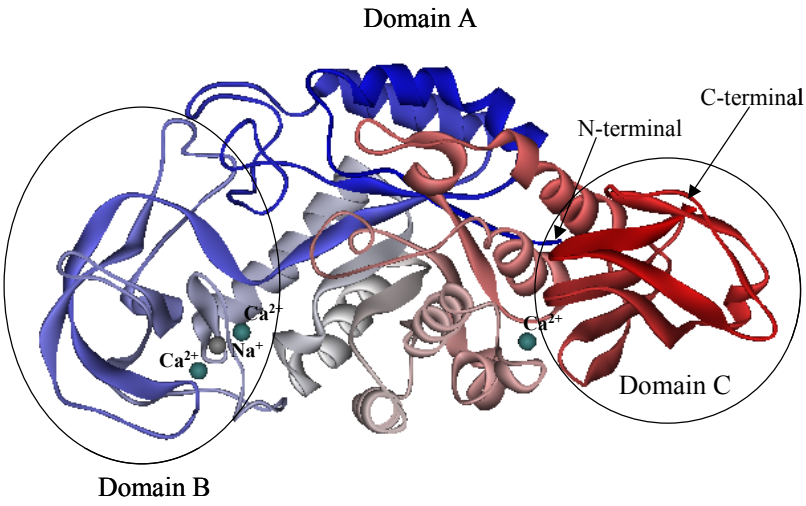

Figure 1.3 Tertiary structure of the wild type *Bacillus halmapalus* α-amylase (BHA). The $(\alpha/\beta)_8$-barrel constitutes domain A. Domain B, which consists of an extended loop, is inserted between β-3 and α-3 of the $(\alpha/\beta)_8$-barrel. Domain C is formed by a C-terminal eight stranded β-sheet domain (Davies et al., 2005; Lyhne-Iversen, 2005).

The retaining glycoside hydrolysis is catalysed via a two-step reaction mechanism which requires the presence of two carboxyl containing amino acids: the first acts as an acid/base catalyst, and the other as a nucleophile forming the glycosyl-enzyme intermediate. The catalytic residues of α-amylases are considered to consist of two aspartic acids and one glutamic acid. The enzyme activity is severely reduced when one or more of these three amino acids were replaced by others (Payan & Qian, 2003). For the wild type *Bacillus licheniformis* α-amylase the three active site residues are Asp231, Glu261 and Asp328 (Nielsen & Borchert, 2000). In contrast, thermostability is greatly influenced by His133, His235 and Ala209 (Upadek & Kottwitz, 1992). The structural stability of the enzyme should be considered with priority in the design of downstream processes. Precipitating agents such as salts or surfactants should be chosen in compliance with stability properties of the enzyme

to be processed. It is needless to say that processing outside the pH- and temperature range tolerated by the protein is inappropriate. The development of a new or the optimisation of an existing enzyme must therefore be considered as holistic processes which should not only be driven by requirements of the final application but should cover all relevant aspects in production and purification.

1.5.2 Applications of α-amylases

α-Amylases are primarily used for starch liquefaction, textile sizing, bread improvement, pulp and paper production, brewing, alcohol production and as additive in the detergent formulations (van Ee, 1992). The process conditions at which the amylases are employed may vary for each individual application. As an example, for economical and ecological reasons, the temperature typically employed in household washing processes is becoming lower. As the solubility of starch is much higher at higher temperature, the α-amylases have to be more efficient and exhibit their activity maximum at lower temperatures to ensure satisfactory washing performance. Moreover, the use of α-amylases as additives in detergents is very demanding with respect to activity and stability. In many cases the washing process is conducted at very high pH (up to 10.5) and the environment can be very oxidising. In addition, the amylases have to be resistant to surfactants, proteases and metal ion chelating agents which are common compounds present in detergents (Upadek & Kottwitz, 1992). A very different environment is found when amylases are employed in starch liquefaction. The rapid liquefaction of starch is necessary to reduce the viscosity of the starch slurry and takes place together with steam injection. This process is conducted at high temperature (95 to 105°C) at pH 6 and even lower pH-values would be desirable to reduce the formation of by-products (Bisgaard-Frantzen et al., 1999). The very different demands can obviously not be met by one single amylase; the consequentially optimisation of amylases for each individual industrial application has substantially benefited from advances in protein engineering. Nevertheless, the optimisation is still a time-consuming process which is difficult to rationalise (Nielsen & Borchert, 2000).

1.6 Methods and techniques employed during the experimental work of this thesis

To study the crystallisation and solubility behaviour of enzymes, a number of different methods have been used during this work. The principles behind these methods are described in the following chapters.

1.6.1 Dynamic light scattering

Dynamic light scattering (DLS), which is sometimes also referred to as photon correlation spectroscopy or quasi-elastic light scattering, is employed to determine the particle size within the sub-micron range (typically from 1 nm to 1 μm). In DLS, scattering intensity fluctuations are monitored in micro-second scale and then correlated. The intensity fluctuations are a result of particle motion (Brownian movement) and the measured property in the correlation analysis is the distribution of diffusion coefficients. The particles in solution are in a constant random Brownian movement which causes the intensity of the scattered light to fluctuate as a function of time. The detected scattered intensity is then taken to construct the autocorrelation or self-similarity function (function of time) which for monodisperse particles is a single exponential decaying function from which the corresponding diffusion coefficient can be determined. The correlation for a large particle takes longer to decay than for a small particle due to the respectively slower or faster Brownian movement. The size of a particle is calculated from the translational diffusion coefficient by the Stokes-Einstein equation:

$$d(H) = \frac{kT}{3\pi\eta D} \qquad (3)$$

where d(H) is the hydrodynamic radius, k the Boltzmann constant, T the absolute temperature, η the dynamic viscosity and D the translational diffusion coefficient. The diameter so obtained is the diameter of a sphere that has the same translational diffusion coefficient as the particle. This method has the advantage of being concentration independent, only requires small sampling volumes, is non-invasive and could be used as an online control tool as results are obtained in real time. However, the samples to be analysed should be

completely free of dust or air bubbles and should be filtered which may not always be possible in industrial processes.

1.6.2 Static light scattering

In contrast to dynamic light scattering (DLS), static light scattering (SLS) considers time averaged scattering intensities observed at one specific scattering angle. One has to consider that the particles are no point scatterer such that their scattering pattern depends on their shape and size. If primary light waves are scattered on several scattering centres the resulting secondary waves differ in their path lengths. This difference of the path lengths results into a phase factor difference of the scattered light waves. Up to a specific size having the same range as the scattered light wavelength, particles are not point scatterer anymore. The interference of the scattered light is resulting from scattering centres of different particles as well as from scattered intensities for a single large particle in the presence of intraparticle interference. The scattering intensity stemming from a particle is angle dependent and leads to the particle form factor P(q). With increasing concentration of the particulate dispersion one can obtain information about the static structure factor S(q) from the time averaged intensities. The structure factor represents the structural arrangement of the particles and is determined by the particle interactions.

The second osmotic virial coefficient

The second osmotic virial coefficient B_{22} is a unit to characterise weak protein-protein interactions; its use has been stimulated by an astonishing correlation between B_{22} and crystallisation conditions: solution conditions at which proteins have an enhanced propensity to crystallise correspond to slightly negative B_{22}-values, resulting from weak attractive protein interactions. The range of slightly negative B_{22}-values of between -0.5 and -8.0×10^{-4} mol mL g^{-2} is thus termed the crystallisation slot and has been proven to be valid for a number of proteins (George & Wilson, 1994). B_{22} can be determined by a number of methods such as X-ray scattering (Ducruix et al., 1996), neutron scattering (Velev et al., 1998), membrane osmometry (Haynes et al., 1992) and sedimentation equilibrium (Behlke & Ristau, 1999). Probably the most popular method is by static light scattering (Georgalis & Saenger, 1999). Here B_{22} is defined as

$$\frac{KC_P}{R_\theta} = \frac{1}{M_W} + 2B_{22}C_P \qquad (4)$$

where M_W is the molecular weight of the protein and C_P is the protein concentration. The Raleigh ratio R_θ is the normal scattered intensity (Zimm, 1948) at given angles. K is a constant calculated from the optical properties of the system as

$$K = \frac{4\pi^2 {n_0}^2}{N_A \lambda^4}\left(\frac{dn}{dC}\right)^2 \qquad (5)$$

where n_0 is the refractive index of the solvent, (dn/dC) is the refractive index increment of the protein, N_A the Avogadro number and λ is the wavelength of the laser in vacuum. These two equations enable to process SLS-data to determine B_{22} (Velev et al., 1998).

Correlation between solubility and second osmotic viral coefficient

What makes the B_{22} particularly interesting for bulk protein production is that according to Haas, Drenth and Wilson (1999) there is a relation between solubility and B_{22} in aqueous protein solutions. This may not be surprising since both B_{22} and solubility are determined by the interactions between protein molecules. This relation is, however, not trivial since the solubility depends on the binding energy between proteins molecules at a short distance in the crystal for very specific orientations of the molecules with respect to each other whereas the B_{22} is a statistical average over all distances and orientations of two molecules in the liquid phase, with each configuration weighted by a Boltzmann average. In spite of these differences, for a large amount of data available for lysozyme, it was found that all B_{22}-solubility-pairs fall approximately on a single curve (Guo et al., 1999). B_{22} can be linked to the solubility S according to

$$B_{22} = \frac{4}{M\rho}\left[1 - A\left\{\left(\frac{S}{m}\right)^{-\frac{2}{z}} - 1\right\}\right] \qquad (6)$$

Here, M is the molecular weight of the protein, ρ is the protein density and

$$m = \frac{M}{18\rho} \tag{7}$$

The Haas-Drenth-Wilson-model (HDW) consists of only two parameters which have to be adjusted; the first one is z which is the coordination number and presents the number of nearest neighbouring protein molecules inside a crystal lattice and usually depends on the crystal structure and the packing fraction. Alternatively, z can be interpreted as the number of macro-bonds in the crystal lattice (Demoruelle et al., 2002). The other parameter A is a constant and is characteristic of each individual protein (Haas et al., 1999). One of the appealing features of predicting the solubility by B_{22} is that the latter can be determined from dilute protein solutions distinctly below the solubility limit such that labour-intensive solubility experiments can be circumvented which require high amounts of crystallisable protein which is often not given. Since B_{22} can essentially be measured in real-time, this correlation would enable the identification and circumvention of process conditions in which the protein is likely to precipitate, thus avoiding unintentional interruption of the product flow and consequent delays in delivery.

However, according to the Rayleigh approximation, the scattering intensity I is highly proportional to the diameter d of the scatterer as $I \sim d^6$ which means that the scattered light from larger particles will quickly superimpose the light stemming from smaller ones (Wyatt, 1993). This high dependence of the intensity on the diameter of the scatterer causes real problems to measure meaningful B_{22} values in systems which are not completely free of dust particles or are prone to form soluble aggregates. This constraint could seriously limit the applicability of SLS in industrial solutions (Skouri et al., 1995).

1.6.3 Self-interaction chromatography

An alternative method to determine the B_{22} would be self-interaction chromatography, SIC, (Patro & Przybycien, 1996) where protein is covalently immobilised on chromatographic

resins which are then packed into a column; the retention time of a pulse of protein injected into the column under isocratic conditions is then measured. The relative retention reflects the average protein interactions. Although this method has originally been developed to characterise stabilising or destabilising effects of additives on protein interactions in qualitative terms (Patro & Przybycien, 1996), it is possible to use SIC to determine B_{22} and thereby describe protein interactions in quantitative terms (Tessier et al., 2002a). The interactions between two protein molecules in solution can be described in terms of the B_{22} as (Zimm, 1948)

$$B_{22} = -\frac{1}{2}\int_{\Omega_2}\int_{\Omega_1}\int_0^{\infty}\left(e^{-\frac{W}{kT}}-1\right)r_{12}^2 dr_{12} d\Omega_1 d\Omega_2 \tag{8}$$

The potential of mean force (PMF) W describes the anisotropic interaction energy between two molecules in solution and is a function of all orientations and separation distances (r_{12}). Ω_1 and Ω_2 are normalised vectors describing the angular position and orientation of both molecules; the factor 1/2 corrects for double counting of an identical pair of molecules. The integral in equation can be split into excluded volume and intermolecular contributions:

$$B_{22} = \frac{1}{2}\int_{\Omega_2}\int_{\Omega_1}\left[\frac{1}{3}r_c^3 - \int_{r_C}^{\infty}(e^{-\frac{W}{kT}}-1)r_{12}^2 dr_{12}\right]d\Omega_1 d\Omega_2 \tag{9}$$

where r_C (Ω_1, Ω_2) is the separation distance upon contact. Protein interactions that dominate the value of B_{22} are typically of short range persisting over a distance less than the diameter of the protein molecule (Neal et al., 1998). An additional relevant parameter is the typical pore size of a chromatographic particle which is usually much larger than the protein diameter. The immobilised protein molecules can therefore be considered as being fixed to a flat surface; moreover it is assumed that a free protein molecule interacts with only one immobilised protein molecule at a time (two-body interaction). Therefore, the particle surface coverage of the immobilised protein has a decisive impact on the interactions. Finally, the free protein molecules are assumed to interact only with the immobilised protein and not with each other which can be warranted by using low protein concentrations in the mobile phase. The

experimentally obtained chromatographic retention time is typically given by the retention factor

$$k' = \frac{V_r - V_0}{V_0} \tag{10}$$

where V_r is the retention volume required to elute a solute from the column and V_0 is the retention volume for a situation in which the free molecules do not interact with the surface of the particles. The B_{22} can be written as

$$B_{22} = B_{22}^{HS} - \frac{k'}{\rho_s \phi} \tag{11}$$

where B_{22}^{HS} is the hard sphere or excluded volume contribution, ρ_s the number of immobilised molecules per unit area, and the phase ratio defined as $\phi = A_S/V_0$ where A_S is the total available surface area. Equation (11) connects the B_{22} and the retention factor based only on the size of the molecule, the amount of immobilised protein per unit area and the phase ratio, the latter a characteristic unit of the chromatographic resin employed (DePhillips & Lenhoff, 2000). SIC is probably not as sensitive as SLS to the presence of aggregates and dust particles and may thus be suited to be used as a controlling tool in large scale protein processing.

1.6.4 Particle sizing

Low angle light scattering

Low angle laser light scattering (LALLS) measures the scattered intensity of a particulate solution over time intervals much longer than typically needed for molecular rotation and translation at several angles. The particle size distribution is determined either by the Mie or the Fraunhofer theory. The Fraunhofer theory, which is an approximation of the Mie-theory, is only able to give accurate results for particles larger than 1 µm in size. In contrast, the Mie theory delivers accurate particle measurements within a size range between 0.02 and 2000 µm. The Mie theory assumes a spherical particle shape, can be employed for transparent and

opaque particles and considers primary scattering from the surface of a particle and predicts secondary scattering caused by light refraction within the particle (Jones, 2003).

Electrical sensing zone method (Coulter principle)

The electrical sensing zone method of sizing and counting particles is based on measurable changes in electrical resistance produced by nonconductive particles suspended in an electrolyte. A small opening (aperture) between electrodes is the sensing zone through which suspended particles pass. In the sensing zone each particle displaces its own volume of electrolyte which is measured as a voltage pulse, the height of each pulse being proportional to the volume of the particle. The quantity of suspension drawn through the aperture is precisely controlled to allow the system to count and size particles for an exact reproducible volume. Several thousand particles per second can individually be counted and sized with high accuracy. This method is independent of particle shape, colour and density. Problems arise from interferences caused by primary and secondary coincidence. The former occurs when two or more crystals pass through the orifice in close proximity, so that their signals overlap and might be interpreted by the device as resulting from a single crystal, whereas secondary coincidence occurs when the crystals pass through the orifice almost simultaneously, so that the resulting pulse is the sum of the pulses the crystals would have caused individually. These problems can be minimised by diluting the solution. Porous particles of conductivities close to that of the electrolyte may be ignored or their size underestimated. Another fact to consider is that the pulse induced by the particle is proportional to the volume. Conversions from volume to length distributions typically assume spherical particle shape which may cause errors in case the particle shape is significantly different from a sphere such that corrections may be required (Wynn & Hounslow, 1997). Calibration, e.g., with Latex beads of a very narrow and well defined size distribution, is necessary for each buffer (and temperature, if applicable) to ensure accurate measurements even at low conductivity. A big advantage of the electrical sensing zone method is that it not only provides a size distribution in terms of fractional numbers but counts the particles so that a particle concentration can be measured. This is particularly interesting if a crystallisation process is characterised by the crystal concentration as a function of time which would not be possible using low angle laser light scattering.

1.6.5 Zeta potential

The zeta potential is a physical property which is exhibited by any particle or molecule in suspension. The liquid layer surrounding a molecule consists of two parts, an inner region (Stern layer) where the ions are strongly bound and an outer diffuse region where they are less firmly associated (Lin et al., 2003). Within the diffuse layer there is a boundary inside which the ions and molecules form stable entities. When a particle moves, ions within the slipping plane move with it. Ions beyond the boundary stay with the bulk dispersant. The potential at the boundary of the slipping plane, i.e. the surface of hydrodynamic shear, is the zeta potential (Figure 1.4). The magnitude of the zeta potential gives an indication of the colloidal stability of the system. Systems of large negative or positive zeta potentials tend to repel each other and stay in solution and vice versa. The general threshold between stable and unstable solutions is ± 30 mV (He et al., 2003). Solutions of zeta potentials of a magnitude smaller than 30 mV tend to aggregate. The zeta potential can be derived from the electrophoretic mobility of a molecule. When an electric field is applied across an electrolyte, charged suspended particles are attracted towards the electrode of opposite charge. Viscous forces acting on the particle oppose this movement and only when the two forces are in equilibrium, the particles move with a constant velocity. This velocity, also referred to as electrophoretic mobility, is dependent on the strength of the electric field or voltage gradient, the dielectric constant of the medium and the zeta potential. Electrophoretic mobility U_E and zeta potential z are related to each other via the Henry equation:

$$U_E = \frac{2\varepsilon z f(\kappa a)}{3\eta} \tag{12}$$

where ε is the dielectric constant, η the dynamic viscosity and f(κa) the Henry's function in which κ is the reciprocal Debye length which describes the thickness of the electrical double layer and a is the radius of the molecule. Thus, κa measures the ratio between molecule radius and the thickness of the electrical double layer. Following approximation for the Henry's function are commonly made (Ricq et al., 1998):

κa<1: Hückel-approximation; f(κa) = 1

κa>>1: Smoluchowski approximation; f(κa) = 1.5

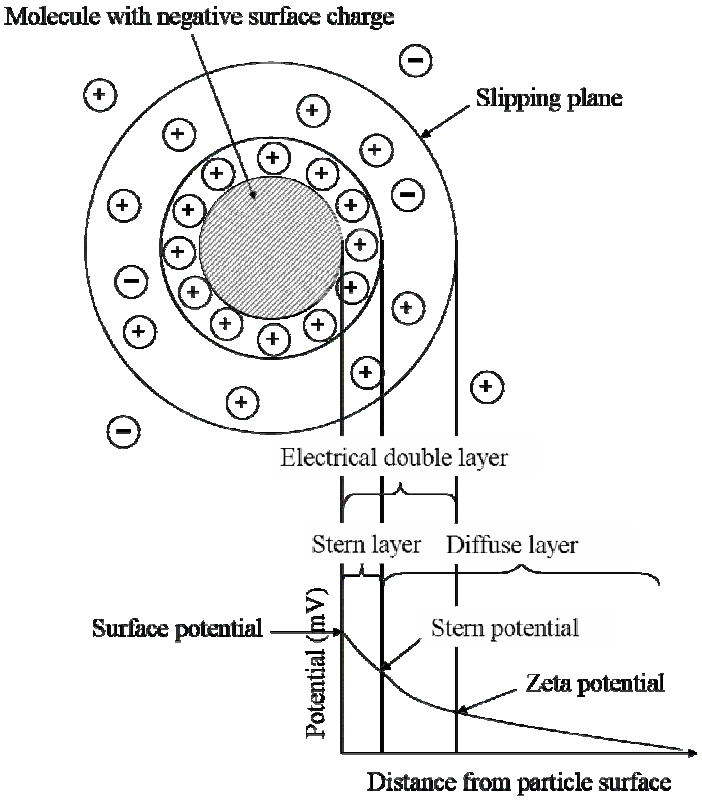

Figure 1.4 **Schematic representation of the ion arrangement surrounding a charged molecule**

When considering a radius of 3 nm which is the range for an α-amylase molecule, and a Debye length of 0.8 nm which is typical for lysozyme at salt-free conditions (Boström et al., 2003; Lee et al., 2001), κa would be 3.75 which means that neither of the two constraints are strictly fulfilled but it appears reasonable to use the Smoluchowski approximation.

The zeta potential e.g., of a protein molecule depends on the pH. The higher the pH, the more negative charges the protein carries and the higher the ability to repel other protein molecules. The zeta potential as a function of pH is therefore correlated to the charge curve; the pH of zero zeta potential equals the isoelectric point of the protein. Moreover, the thickness of the double layer gets compressed with increasing ionic strength (Boström et al., 2003).

The electrophoretic mobility of a system is measured in a cell with electrodes at either end to which a potential is applied. The velocity by which the particles move to the oppositely charged electrode is expressed in unit field strengths as their mobility which is measured by Laser-Doppler-Velocimetry. The optics is focused to relay the scattering of the particles in the cell. The scattered light is measured at an angle of 17° and combined with a reference beam. The rate of intensity fluctuation is proportional to the speed of the particle. A combination of Laser-Doppler-Velocimetry and Phase Analysis Light Scattering enables the measurement of samples of very low mobility such that even protein molecules as small as lysozyme in solution can be characterised by means of the zeta potential (Dai et al., 2004).

1.6.6 Scanning electron microscopy

An electron source produces a stream of monochromatic electrons which is condensed by a first condenser lens which is used to form the beam and to limit the amount of the current in the beam and works in conjunction with a condenser aperture to eliminate the high-angle electrons, i.e. electrons far away from the optic axis of the microscope which may disturb the signal. The second condenser lens forms the electrons into a thin, tight, coherent beam. A user selectable objective aperture further eliminates high-angle electrons from the beam. A set of coils then move the beam in a grid fashion (like a television), dwelling on points for a period of time determined by the scan speed (usually in the microsecond range). The final lens, the objective, focuses the scanning beam onto the desired part of the specimen. The signals which are generated when the beam strikes the sample are detected, counted and translated into a pixel the intensity of which is determined by this number of interactions. The three signals which provide the greatest amount of information in SEM are the secondary electrons, backscattered electrons, and X-rays. This process is repeated until the grid scan is finished and then repeated, the entire pattern can be scanned 30 times per second. The samples which are to be analysed must be stable under vacuum which prevents the examination of fluid-containing matter.

Environmental scanning electron microscopy

The field-emission environmental scanning electron microscope (ESEM-FEG) represents several important advances in scanning electron microscopy. Whereas conventional scanning

electron microscopy requires a relatively high vacuum in the specimen chamber to prevent atmospheric interference with primary or secondary electrons, ESEM may be operated with a poor vacuum (up to 10 Torr of vapour pressure, or one seventy-sixth of an atmosphere) in the specimen chamber. In such wet mode imaging, the chamber is isolated (by valves, pressure-limiting apertures, and a large-diameter bypass tube) from the rest of the vacuum system. As water is the most common imaging gas, a separate vacuum pump permits fine control of its vapour pressure in the specimen chamber. When the electron beam (primary electrons) ejects secondary electrons from the surface of the sample, the secondary electrons collide with water molecules, which in turn function as a cascade amplifier, delivering the secondary electron signal to the positively biased gaseous secondary electron detector (GSED). Because of this electron loss in this exchange, the water molecules are positively ionized, and thus they are forced/attracted toward the sample (which may be nonconductive and uncoated), serving to neutralize the negative charge produced by the primary electron beam. Because of the lower demand on the vacuum, no special sample preparation is needed such that samples can be examined at conditions close to their original environment (McDonald, 1998).

1.6.7 Protein and activity assays

Total protein

Since the protein concentration is the key parameter in crystallisation processes, the determination is of essential importance and calls for careful considerations. Probably the easiest and most commonly employed method is to determine the protein concentration by UV-absorbance at a wavelength of 280 nm at which primarily the two amino acids tyrosine and tryptophan absorb (Judge et al., 1996; Pusey & Munson, 1991). This is probably a well suited method for very pure protein solutions but may cause problems in the presence of nucleic acids which absorb within the same range of wavelengths. To enhance the comparability with other protein solubility data, an ESL-protein assay purchased from Roche, Mannheim, Germany, was employed in this study. In the assay, reaction of Cu^{2+} ions and NaOH with protein in the sample first results in the formation of a Biuret compound (i.e. a Cu(II)-protein chelate complex). In the second step, excess Cu^{2+} ions are reduced to Cu^{+} by ascorbic acid. The Cu^{+} so produced is then chelated by 2,9 Dimethyl-4,7-diphenyl-1,10-phenanthroline-disulfonate (bathocuproine disulphonate) to form a Cu(I)-bathocuproine

complex. The amount of Cu(I)-bathocuproine complex formed is inversely proportional to the to the amount of peptide bonds and its absorbance can be read at 485 nm wavelength. It is not influenced by specific amino acid side chains so that the protein-protein variability is reduced (Matsushita et al., 1993). Bovine serum albumin was used to determine the standard curve.

Enzyme activity

The α-amylase activity during crystallisation experiments was determined by measuring the absorbance of p-nitrophenol (PNP) at a wavelength of 405 nm (Lorentz, 2000). PNP is formed by the α-amylase catalysed degradation of a blocked ethylidene-G7-PNP substrate (G: glucose). A constant specific activity (i.e. α-amylase activity divided by the protein concentration) was taken as evidence that the protein could maintain its structural integrity during experimentation.

1.7 Aims and scopes of the thesis

The overall aim of this thesis was to evaluate traditional and novel approaches for the characterisation and quantification of the influence of different solution properties such as changes in salt concentration, pH and temperature on protein solubility, crystal growth and the overall phase diagram. The tools were evaluated in view of the applicability under process relevant conditions with the aim of reducing protein demand whilst operating in the presence of impurities.

The specific aims were to use conventional batch crystallisation approaches to extensively describe the solubility of a recombinant *Bacillus halmapalus* α-amylase (BHA) at different temperatures, pH and selected anions and cations from the Hofmeister series (Chapter 2). The kinetics of the batch crystallisation process of BHA were described in terms of the development of the crystal size distribution, the supernatant protein concentration and the crystal concentration as a function of time (Chapter 3). The validity of the hypothesis that the Hofmeister series is reversed depending on the sign of the net charge was reviewed by studying the solubility of *Bacillus licheniformis* α-amylase (BLA) in the presence of different salts on both sides of the isoelectric point (Chapter 4). The applicability of 96-well microtitre

plates to generate complete phase diagrams consisting of precipitation, nucleation, metastable and undersaturated zones is described in Chapter 5. The potential of self-interaction chromatography as an alternative method to static light scattering to determine the second osmotic virial coefficient B_{22} and to correlate it to solubility was examined in Chapter 6. In the final chapter the different approaches chosen in this thesis are put into perspective and a strategy is proposed how they can help developing reliable, flexible and efficient downstream processes tailored for novel proteins.

2 Factors affecting the solubility of *Bacillus halmapalus* α-amylase

Cornelius Faber,[a,b] Timothy J. Hobley,[b,*] Jørgen Mollerup,[c] Owen R. T. Thomas,[b,d] and Svend G. Kaasgaard,[a]

[a] Novozymes A/S, Novo Alle, DK-2880 Bagsværd, Denmark

[b] Center for Microbial Biotechnology, BioCentrum-DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[c] Department of Chemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[d] Department of Chemical Engineering, School of Engineering, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

*Corresponding author

e-mail: th@biocentrum.dtu.dk (T. Hobley)

Telephone: +45 45 25 27 06

Fax: +45 45 88 41 48

Synopsis: The influence of pH, temperature and selected salts from the Hofmeister series on the solubility of *Bacillus halmapalus* α-amylase is studied and discussed.

Keywords: anions; cations; crystallisation; Hofmeister series; polydispersity; proteins; retrograde; salts; zeta potential;

2.1 Abstract

A detailed study of the solubility of recombinant *Bacillus halmapalus* α-amylase has been conducted. A semi-purified preparation from a bulk crystallisation was chosen that contained six isoforms with pH-values of between 5.5 and 6.1. The solubility was strongly affected by pH and could be reduced approximately 200-fold at pH 6 as compared to pH 10, leaving only 0.1 mg/mL in solution. Solubility could also be dramatically manipulated using salts. The choice of anions was found to be more important than of the cations, and the lowest solubility was found using sodium sulphate. For the anions, solubility followed the order expected from the Hofmeister series, however, a more complex behaviour was seen for the cations. With the exception of lithium their efficiency to influence the solubility was reversed to what was expected. The polydispersity of the solution was reduced by salt addition and zeta potential measurements indicated a shift in pI caused by lithium. Possible explanations for the observations are discussed, extending our present understanding of how salts affect the solubility of proteins, one that to date is primarily based on experiments with lysozyme.

2.2 Introduction

For many materials, including organic compounds such as antibiotics, organic acids and amino acids, bulk crystallisation has long been an important, cheap and scalable purification and separation process. However, the crystallisation of biological macromolecules is more difficult than for many inorganic or small organic molecules, and with few exceptions, is only just starting to emerge as a versatile operation for large-scale protein recovery. One of the most important examples of industrial-scale crystallisation for protein recovery and purification is the production of insulin (Brange et al., 1988). The crystallisation of some bulk enzymes also plays an important role in their purification and concentration (Jacobsen et al., 1998). Due to continued optimisation of the strains used for production of bulk enzymes, improved fermentation methods and demands for even higher product strengths, there is considerable risk of operating above the solubility limit, even when crystallisation is not desired. For industrial processes it is thus very important to know which factors control the solubility and crystallisation behaviour of the proteins. However, protein crystallisation

literature is primarily orientated towards the use of micro-crystallisation techniques, where the goal is to obtain a few crystals suitable for structure analysis by X-ray diffraction.

When operating above the solubility limit in the large scale production of proteins, e.g., in a crystallisation process, the main focus is on: solubility; the formation of crystals of uniform size, which makes them easy to harvest; and the robustness of the process. The control of nucleation, crystal growth, and a comprehensive knowledge of solubility properties under relevant conditions are of key importance here. The availability of detailed solubility data, or at the least information on how various precipitants (e.g., presence of certain salts) affect the solubility phase diagrams, would lead to better prediction of the crystallisation behaviour and, thus to better process control. However, comprehensive solubility data is only available for very few proteins, most notably lysozyme of high purity. The crystallisation of ovalbumin (Judge et al., 1995), and the nucleation and the growth of microbial lipase from clarified concentrated fermentation broths (Jacobsen et al., 1998) are among the few examples of scientific work published on bulk crystallisation in the presence of impurities. Although the knowledge of solubility properties in industrial processes was regarded as important in both works, only little data on solubility was provided. Furthermore, whilst Hofmeister (Hofmeister, 1888) proposed a series ranging the ability of anions and cations to precipitate proteins, considerable discrepancies have been reported for different proteins. Anions were ranked for their ability to precipitate hen egg white proteins as: sulphate > hydrogen phosphate > acetate ≈ citrate > tartrate > hydrogen carbonate > chromate > chloride > nitrate > chlorate > thiocyanate. It has been found that in general the Hofmeister series is followed in the order listed above by acidic proteins such as *Hypoderma lineatum* collagenase, whereas it is reversed for the basic proteins such as lysozyme (Riès-Kautt & Ducruix, 1997).

Next to anions, the study presented here aims to contribute to the knowledge on the influence of cations on the solubility of acidic proteins. It is widely accepted that two main effects govern the solubility behaviour of proteins, namely electrostatic and hydration interactions (Melander & Horvath, 1977; Petsev & Vekilov, 2000). Electrostatic interactions follow the same principles whether anions or cations are involved, only the sign and extent vary depending on the protein's net charge and the type of ions. Hydration interactions, however, reveal significant differences, depending on whether cations or anions are involved and are believed to increase with the ion concentrations but should be independent of the protein's net charge (Collins & Washabaugh, 1985). A consequence would be that the effect of cations on

the solubility would also invert depending on the sign of the protein's net charge within a concentration range in which electrostatic dominate over hydration interactions.

In the Hofmeister series, cations are ranked according to their hydration entropy. Small singly charged cations such as Li^+ are strongly hydrated, whereas large singly charged cations like Cs^+ are weakly hydrated (Bénas et al., 2002). Li^+ carries a higher surface charge density and therefore a more negative hydration entropy than e.g., Cs^+. The following series for monovalent cations is thus obtained: $Cs^+ > Rb^+ > K^+ > NH_4^+ > Na^+ > Li^+$. This ranking is the reverse of what is commonly referred to as the correct order of the Hofmeister series for monovalent cations found for lysozyme by crystallisation studies. Two other series for cations have been established according to how they affect solubility. The first one by Green (Green, 1932) was carried out on carboxyhemoglobin near its pI of around 6.6 and led to following ranking: $Li^+ > Na^+ \sim K^+ > NH_4^+ > Mg^{2+}$. The second one was established by Bénas et al. (Bénas et al., 2002) on lysozyme at pH 4.5 and 25°C and resulted in following ranking: $Rb^+ > Cs^+ > Co^{2+} > Mn^{2+} > Yb^{3+}$

Here, we present the results of a detailed study on the solubility in aqueous solution of recombinant α-amylase from *Bacillus halmapalus* (BHA). This model protein was chosen because it readily forms crystals and exhibits a wide solubility window under selected conditions.

2.3 Materials and methods

2.3.1 Preparation of *Bacillus halmapalus* α-amylase

Recombinant *Bacillus halmapalus* α-amylase was expressed in *Bacillus licheniformis* grown on a complex medium, and purified at Novozymes A/S, Bagsværd, Denmark. Cells and other solids were removed by filtration, and the resulting solution was concentrated by ultrafiltration (10 kDa cut off), before crystallising the enzyme by lowering the pH to 7.5. The crystals were then harvested by centrifugation.

2.3.2 Reagents

Sodium hydroxide and acetic acid (at concentrations of 0.1 and 1 M respectively) and potassium nitrate were obtained from Bie & Berntsen (Rødovre, Denmark) and sodium acetate and sodium nitrate were purchased from Merck (Darmstadt, Germany). Sodium chloride was acquired from Fluka (Buchs, Switzerland), whereas sodium sulphate and calcium chloride were provided by J.T. Baker (Deventer, The Netherlands) and sodium thiocyanate, lithium nitrate, caesium nitrate and rubidium nitrate by Sigma-Aldrich (Steinheim, Germany). All chemicals were of analytical grade.

2.3.3 Sample preparation

For all experiments, a buffer containing 10 mM MES, 10 mM HEPES (both from Sigma-Aldrich) and 10 mM boric acid (AppliChem, Darmstadt, Germany) was used. These reagents were dissolved in ultra-pure water of 18.2 MΩcm resistivity (Millipore, Billerica, MA, USA) and 0.2 % (w/v) of the antimicrobial agent, Proxel LV (Avecia, Manchester, UK), was added. Crystals of α-amylase were then dissolved in the buffer at pH 10.5, and after dissolution, the solution was filtered through a 0.22 μm pore size cellulose acetate filter (Sartorius, Göttingen, Germany).

2.3.4 Crystallisation studies

All crystallisation experiments were conducted in batch mode. For crystallisation under every condition tested, solutions of two different initial protein concentrations were prepared in the buffer described above and containing added precipitants (i.e. salts) as needed. For each preparation, the crystallisation was then initialised by slowly (over 5 min) reducing the pH to the value of interest. The solution was immediately aliquoted into a series of 1.5 mL Eppendorf tubes (750 μL sample volume), which were sealed and placed on a thermo mixer (Model 5355, Eppendorf, Hamburg, Germany) that permitted precise temperature control (± 1°C) and suitable mixing conditions (1400 rpm). No crystals or precipitates were observed during the above procedure. At each given experimental time, two of the tubes from each series were removed from the mixer and the formation of crystals was examined by light

microscopy. The crystals were then pelleted at 25,000 g_{av} for 2 minutes in a microcentrifuge (Model 5415 R, Eppendorf). The temperature employed during centrifugation was the same as that used during crystallisation. The supernatants were collected and filtered through 0.22 μm filters (Sartorius) and kept at -30°C until analysis. Upon thawing, no crystals were observed in any samples, all of which passed through no more than two freeze-thaw cycles (the enzyme's activity was unaffected by this treatment). The pH was checked at the end of a given crystallisation process, and no adjustments were made during an experiment.

2.3.5 Analysis

Supernatants were analysed for total protein by a reverse biuret method combined with copper-bathcuproine chelate reaction (Matsushita et al., 1993). The assay was scaled for use in a Cobas Fara spectrophotometric robot (Roche, Rotkreutz, Switzerland). The absorbance was measured at 485 nm and protein concentrations were expressed as bovine serum albumin equivalents. Enzyme activity was quantified by the AMYL kit (Roche, Mannheim, Germany), which was scaled for automation on an ELISA plate-reader system (Spectra max plus, Molecular Devices, Sunnyvale, CA, USA). The assay is based on the formation of *p*-nitrophenol (PNP) upon α-amylase catalysed cleavage of a blocked PNP-glucose substrate, which can be quantified spectrophotometrically at 405 nm (Lorentz, 2000). In the current work, all the specific activities are given as arbitrary units per mg protein normalised to the same value, which is defined as 100 % activity. The salts used in the crystallisation studies were shown not to exhibit any influence on the assays.

Re-dissolved crystals were checked for purity by reducing SDS-PAGE (NuPage 10% Bis-Tris-gel NP0301, Invitrogen, Carlsbad, CA, USA). A mixture of defined proteins of known molecular weight (LMW 17-0446-01, Amersham Biosciences, Uppsala, Sweden) was used to calibrate the SDS-polyacrylamide gel electrophoretograms. The pI was determined by isoelectric focusing (IEF) in pH 3.5 to 9.5 ampholine PAGplate polyacrylamide gels (Product number 80-1124-80, Amersham Biosciences, Uppsala, Sweden), and a commercially obtained protein mixture (Product number 17-0471-01, Amersham Biosciences) was used to calibrate the gels. Both gels from SDS-PAGE and IEF runs were stained with Coomassie Brilliant Blue R-250 (Sigma-Aldrich, Steinheim, Germany). To demonstrate the presence of active amylase in the IEF gels, the bands from an unstained gel were excised, using a stained gel as template.

The gel fragments were placed in 600 μL Eppendorf vials that had been perforated at their base with a fine gauge syringe needle. These vials were then inserted into 1.5 mL Eppendorf tubes and the gel fragments were homogenised into a slurry by passage through the holes under the influence of centrifugal force (i.e. spinning in a microfuge). The 600 μL vials were removed, and homogenised gel slurries, now contained in the 1.5 mL Eppendorf tubes, were suspended with 1 mL of 15 mM $CaCl_2$. After one hour, the samples were centrifuged to remove the leached gel pieces, and the presence of amylase activity in the supernatants was demonstrated qualitatively using the Phadebas kit (Product number 10-5380-33, Amersham Biosciences) following the instructions of the manufacturer.

The polydispersity and the zeta potential of the supernatants obtained from batch crystallisations were measured in a Zetasizer Nano ZS (Malvern Instruments, Malvern, UK). Samples were filtered through 0.22 μm filters (Sartorius). A 1.5 mL aliquot of the diluted sample was transferred into a micro cuvette (Ratiolab, Dreieich, Germany) previously cleaned with MilliQ-water to remove dust particles. The cuvettes were then closed with a lid to avoid dust intrusion. The polydispersity was determined by dynamic light scattering. The instrument was equipped with one detector placed at an angle of 173°. Each result was the average of 30 measurements and each individual experiment was conducted in triplicate. The zeta potential was measured using folded capillary cells of 0.75 mL sample volume provided by Malvern. Phase analysis light scattering (PALS) was used to determine the electrophoretic mobility from which the zeta potential was calculated by the Smoluchowski approximation. Since the protein (< 10 mg/mL) and salt concentrations (0.2 M) were low, the viscosity of the samples was approximated by assuming the viscosity of water. Measurements were conducted in triplicate, each of which consisted of 30 individual measurements. The standard deviation of the mean zeta potential values of the 30 individual measurements was in the range of ± 1 mV.

2.4 Results and Discussion

2.4.1 Properties of the feedstock and crystallisation behaviour

The α-amylase containing feedstock was examined for purity by SDS-PAGE and only one band at the expected molecular weight of 55 kDa was visible (Figure 2.1 A). However, when

isoelectric focusing was used six bands could be observed in a range between pH 5.5 and 6.1 (Figure 2.1 B). In order to test which of these contained amylase activity, four fragments (*a* to *d*) covering this pH range were excised from a gel run in parallel, but not stained (the approximate position of each cut is shown in Figure 2.1 B). After extraction of the proteins from the gel, amylase activity could be detected in all four fragments, indicating that the enzyme is present in various isoforms. No activity could be detected in gel fragment '*e*', in which no protein bands were observed, and thus represents a gel blank.

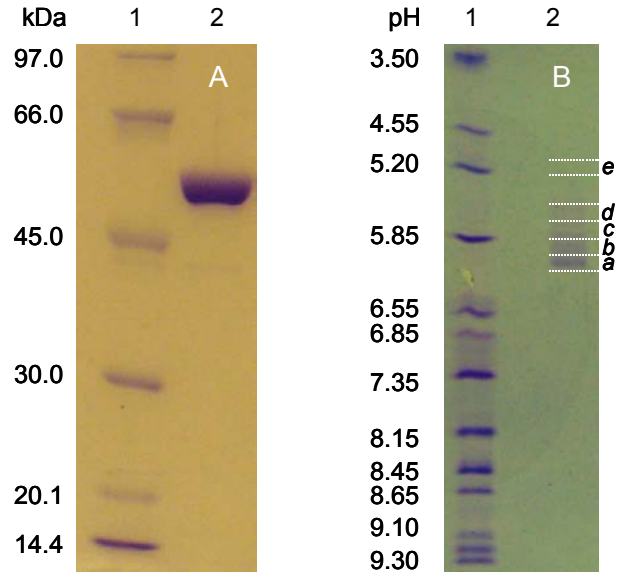

Figure 2.1 Examination of *B. halmapalus* α-amylase feedstock. (A) SDS-PAGE analysis. The molecular weight calibration standards are shown in lane 1. (B) Isoelectric focusing. The calibration standards are in lane 1. A second gel was run in parallel but without staining and the sections marked '*a*–*e*' were excised and analysed for amylase activity.

Consequently, it is concluded that the amylase feedstock was purified to homogeneity as determined by SDS-PAGE, but it is not known how the isoforms differ from each other. Moreover, BHA-solutions exhibited an amber/yellowish colour which indicated the presence of impurities other than proteins, such as sugars and salts.

Prior to studying the effect of temperature, pH or salt on α-amylase solubility, the time required to reach equilibrium was determined. Two different initial protein concentrations were chosen, and supersaturation was then induced by lowering the pH from 10.5 to 9. The first crystals became visible under the microscope after two hours and A shows the time taken

to reach equilibrium for the crystallisation process at pH 9 and 40°C. Within the first 24 h, the protein concentration in the supernatant was significantly lowered, and in the following 72 h there was only a small further reduction. The solution which started from a lower initial supersaturation contained a higher protein concentration in the supernatant after 24 h, but after 96 h the protein concentrations in the mother liquor were almost identical for both experiments, indicating that equilibrium had been reached. When equilibrium was examined under conditions giving one of the lowest solubilities (pH 6 and 5°C) a similar behaviour was observed (B). Due to the significantly reduced solubility in the latter experiment compared to that performed at pH 9 and 40°C, the crystallisation process was started from a lower initial protein concentration, however the relative supersaturation was within a similar range. Again, 96 h was sufficient to reach equilibrium. Although all the solubilities shown subsequently in this work only represent values after 96 h of incubation, samples were also analysed after 24, 48 and 72 h, to ensure that equilibrium was reached under all conditions.

The specific activity of the enzyme measured in the supernatant during crystallisation under each individual condition was essentially independent of time after 24 h at any given condition (Figure 2.2 A and B). The preservation of the specific activity indicated that the enzyme was not significantly changed throughout the experimental time. This is important as a decrease in activity during the experiments would suggest that enzyme modifications had occurred and would be expected to influence solubility.

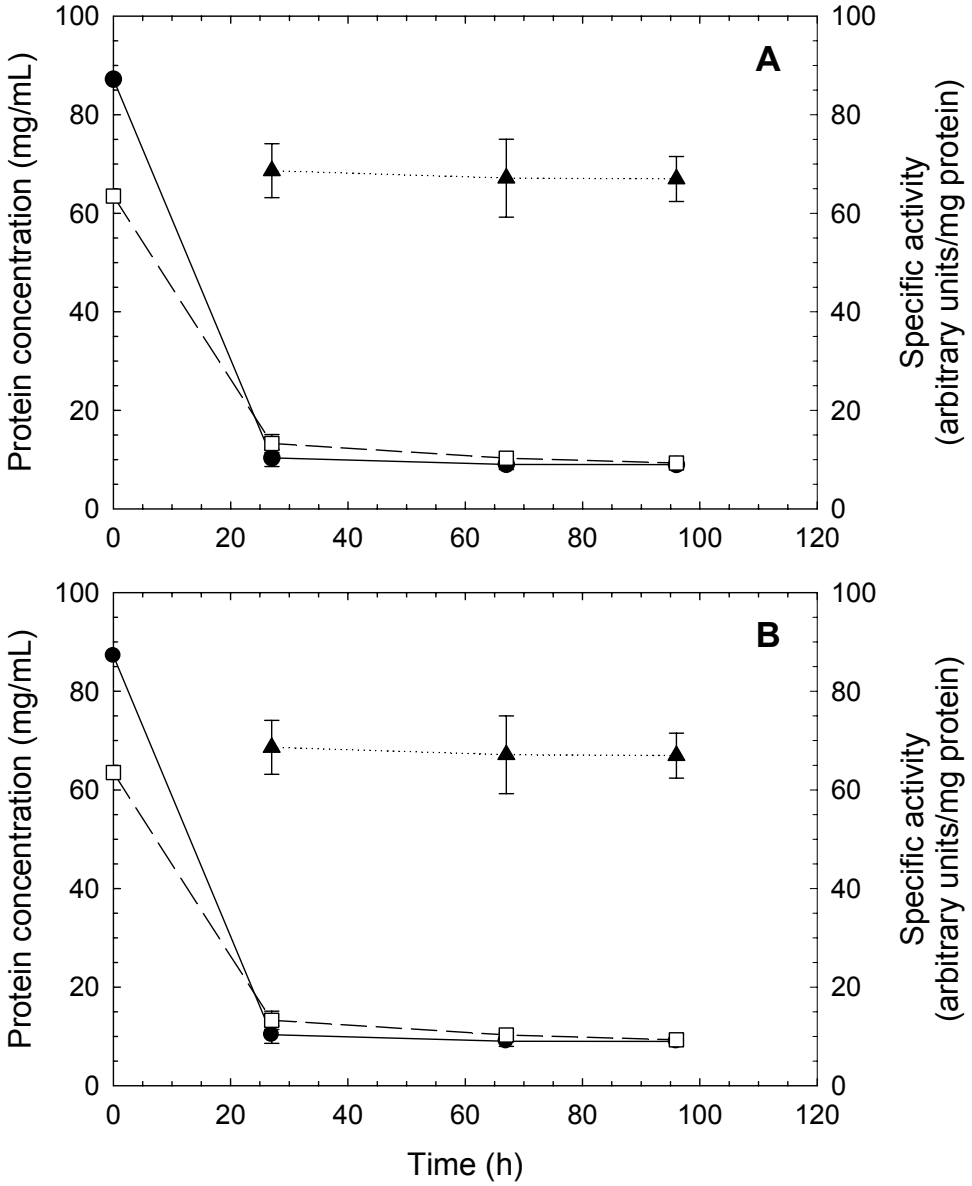

Figure 2.2. α-Amylase solubility and specific activity as a function of time. Protein concentration during crystallisation at (A) pH 9 and 40°C, and (B) pH 6 and 5°C starting from a high (●, solid line) and a low (□, dashed line) initial supersaturation, and the specific enzyme activity (▲, dotted line) averaged for both supersaturations (each supersaturation measured in duplicates) at each given condition. The specific activity was calculated by dividing the average activity obtained from duplicates by the average of the corresponding protein concentration.

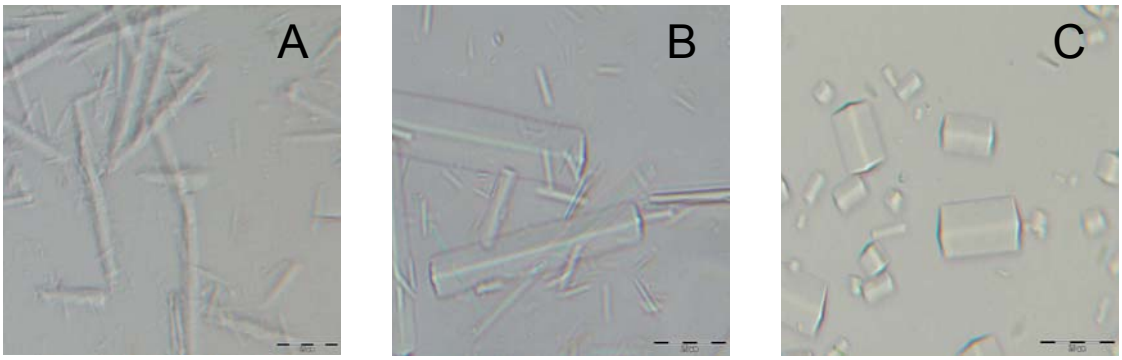

Figure 2.3 Influence of pH on crystal morphologies at 32°C. α-Amylase crystals grown at: (A) pH 6; (B) pH 8 and (C) pH 9.7. The scale bars represent a length of 20 µm.

All experiments were carried out in MHB-buffer (i.e. MES, HEPES, boric acid) to reduce the drift in pH during crystallisation. As the concentration of the buffer salts (10 mM of each) was much lower than the concentration of the salt precipitants (0.1 – 1.0 M), their influence on the solubility is believed to be low. Using this set-up, the shifts in pH during the crystallisation experiments were small. A drift in pH of more than ± 0.2 units after 96 h was thus only observed for experiments conducted at salt concentrations of 0.5 M or higher. These cases are noted where applicable in the results.

2.4.2 Solubility as a function of pH

An inspection of the literature indicates that there have been few studies with proteins in which a large range of pH values have been used to examine solubility. This is because most proteins are not stable over a large pH range. However, it is known that pH changes can lead to crystals having different space groups and morphologies (Riès-Kautt & Ducruix, 1989). The latter is particularly important since morphology affects solubility (Ninomiya et al., 2001) and from a bioprocess viewpoint is critical for efficient crystal harvesting. As mentioned above, the *Bacillus halmapalus* α-amylase was stable over time at pH values between pH 6 and 10 (Figure 2.2). At pH 6, the enzyme crystallises as long, thin, needle-like rods, whereas at pH 8 the rods became broader (~15 µm), but were of similar length (~50 µm), while at pH

9.7, short broad rods were observed (Figure 2.3 A - C). Although the habit of the crystals grown at pH 6 differed somewhat from the ones at pH 8 and 9.7, it is assumed that they were of the same morphology. In the early stage of the crystallisation process, the crystals were of similar shape and size under all conditions. Crystals of the same enzymes as used in this study grown at pH 8 in 25 mM HEPES and 15 mM $CaCl_2$ revealed the same habit as the ones found here (i.e. similar to those in Figure 2.3 B and C) and were determined to belong to space group P222 (Nielsen, 2003). Crystals of another *Bacillus halmapalus* α-amylase variant grown at pH 7.5 and in 50 mM imidazole and 5 mM $CaCl_2$ in hanging drop mode belonged to space group $P2_12_12_1$ (Davies et al., 2005). Two porcine pancreatic α-amylase isoenzymes grown at pH 8, 18°C and in 10 mM Tris-HCl exhibited quite a different crystal habit, but still belonged to the same space group, i.e. $P2_12_12_1$ (Boistelle & Astier, 1992). As no X-ray analyses of the crystals obtained in this study were conducted, no information on their space group can be provided.

The solubility of α-amylase was found to be highly dependent on pH and could be reduced 200-fold by lowering the pH from 9.7 to 6. Under the conditions examined the lowest solubility was at pH 6, which is within the determined pI-range of between pH 5.5 and 6.1 (Figure 2.4). Interestingly, the specific activity decreased with decreasing pH (Figure 2.4), but stayed constant with time under each condition (Figure 2.2). Initially it was suspected that the change in specific activity may be due to the absence of calcium ions in the buffers used since the requirement of calcium ions for activity and structural stability is a characteristic feature of many α-amylases (Declerck et al., 1995; Suvd et al., 2001). It has been shown that addition of excess EDTA to *Bacillus subtilis* α-amylase at 37°C resulted in a 40% loss in activity after 2 h and complete reactivation was not possible upon addition of excess calcium ions (Vallee et al., 1959). Calorimetric studies on *Bacillus halmapalus* α-amylase have demonstrated that thermal inactivation is irreversible, regardless of whether calcium is present or not (Nielsen et al., 2003). However, calcium exerted a strong influence on the denaturation temperature. Nielsen et al. (2003) found that calcium depleted α-amylase first denatured at 48°C, whereas excess calcium increased the denaturation temperature to 83°C. The crystallisation experiments in the present work were conducted without added calcium ions, but as they were carried out at 40°C or lower, i.e. below the denaturation temperature of the calcium-depleted α-amylase, inactivation due to calcium deficiency is not likely. Also the calcium ions are

tightly bound to the enzyme and it generally takes a large excess of EDTA to strip the amylase for calcium.

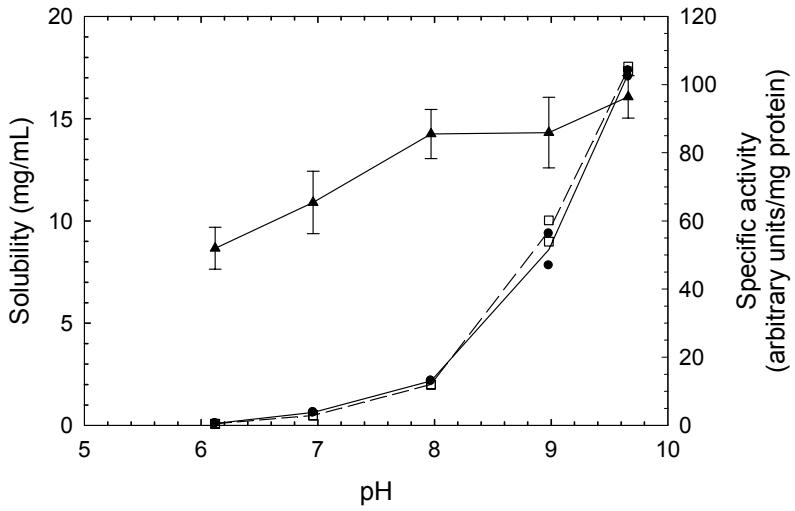

Figure 2.4 The effect of pH on the solubility and specific activity (▲) of α-amylase. Crystallisation experiments were carried out at 32°C starting from higher (●, solid line) and lower (□, dashed line) initial supersaturations. Protein concentrations and activities were measured after 96 h and each experiment was conducted in duplicate. The specific activity was calculated by dividing the average of the four activity measurements (two for each initial supersaturation) by the average of the corresponding protein concentration.

To further confirm that calcium depletion was not the cause of the observed loss of specific activity, the amylase was first dissolved at pH 9, the pH was then lowered to 7, and then increased back to 9 after 24 h. At each step samples were withdrawn for analysis of enzyme activity. The results showed that at pH 7, 76% of the original specific activity was measured. When the specific activity was measured 24 hours after re-adjusting the pH back to 9, a value of 81% or 83% of the original was found in the absence or presence (15 mM) of calcium ions, respectively. As all activity assays were run at the same pH (i.e. 7.5) and in the presence of 40 mM calcium chloride as stabiliser, it is speculated that the change of specific activity as a function of pH may not be caused by calcium depletion, but rather by a rapid and irreversible conformational change. Interference from a protease is less likely, as the specific activity remained constant during the crystallisation.

2.4.3 Solubility as a function of temperature

When the effect of temperature on solubility was studied for *Bacillus halmapalus* α-amylase, a clear retrograde temperature dependence was seen at pH 8.8 between 5°C and 15°C (Figure 2.5). At higher temperatures however, the solubility stayed almost constant. At pH values of 6, 7 and 8, the effect of temperature was much less pronounced than at pH 8.8 (Figure 5). Interestingly, at temperatures higher than those used in this study, Nielsen (Nielsen, 2003) reported that the solubility of the same α-amylase showed a direct temperature dependence at pH 8 and 0.025 M HEPES. A retrograde temperature dependence on solubility has been reported for other proteins and can be much more pronounced. Horse serum albumin exhibited 5 to 10-fold solubility changes even within smaller ranges (0°C to 20°C) (Rosenberger et al., 1993), whereas a strong direct temperature dependence was reported for lysozyme, which changed 5 to 10-fold between 10°C and 35°C. In the case where dissolution processes are exothermic, the solubility decreases with raised temperature (i.e. retrograde dependence) and vice versa (i.e. direct dependence) (Vuolanto et al., 2003). However, the solubility of the α-amylase examined here showed a rather complex temperature dependence, and thus further studies would be required to characterise the thermodynamics of dissolution. The data (Figure 2.5) suggests furthermore that the degree of the temperature dependence also varies with pH. This type of phenomenon has also been reported for lysozyme (Cacioppo & Pusey, 1991b) and xylose isomerase (Vuolanto et al., 2003).

2.4.4 Solubility as a function of salt concentration

Effect of anions

It has been known for more than a century that protein solubility is affected by different concentrations and types of salt (Hofmeister, 1888), and on this basis anions were ranked for their ability to precipitate hen egg white proteins as:
sulphate > hydrogen phosphate > acetate ≈ citrate > tartrate > hydrogen carbonate > chromate > chloride > nitrate > chlorate > thiocyanate

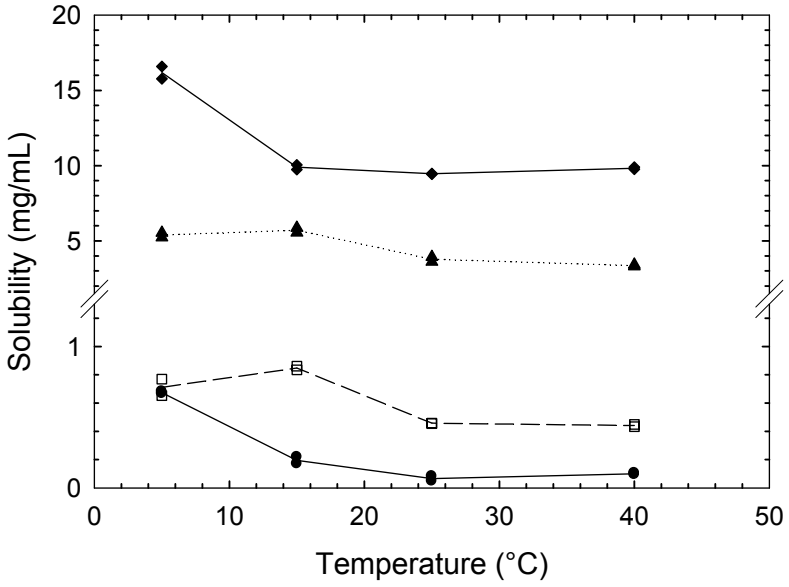

Figure 2.5 Temperature dependence on the solubility at different pH values. Each data point represents the average and the standard deviation of the protein concentrations after 96 h obtained from crystallisation experiments started from two different initial supersaturations. Experiments were carried out at pH 8.8 (◆, solid line), 8.0 (▲, dotted line), 7.0 (□, dashed line) and 6.0 (●, solid line). Each experiment was conducted in duplicate.

It has been found that in general the Hofmeister series is followed in the order listed above by acidic proteins such as *Hypoderma lineatum* collagenase, whereas it is reversed for the basic proteins such as lysozyme (Riès-Kautt & Ducruix, 1997). To shed light on the solubility behaviour of *B. halmapalus* α-amylase, the influence of the sodium salts of acetate, chloride, nitrate, sulphate and thiocyanate at a concentration range from 0 to 1 M was tested (Figure 2.6). At all the examined salt concentrations, the Hofmeister series was followed in the listed order from sulphate, which drastically lowered the solubility, to thiocyanate containing solutions, which resulted in the highest solubility of the amylase. With the exception of thiocyanate, the biggest effects on solubility were found at salt concentrations between 0 and 0.2 M. Further increases in the salt concentration led to only minor reductions in solubility in the case of sodium sulphate and sodium acetate (Figure 2.6). However for the nitrate and thiocyanate salts, a minimum solubility was observed at concentrations between 0.1 and 0.2 M, above which the solubility began to increase. Solubility could in fact be doubled by using 0.5 M thiocyanate, as compared to no added salts.

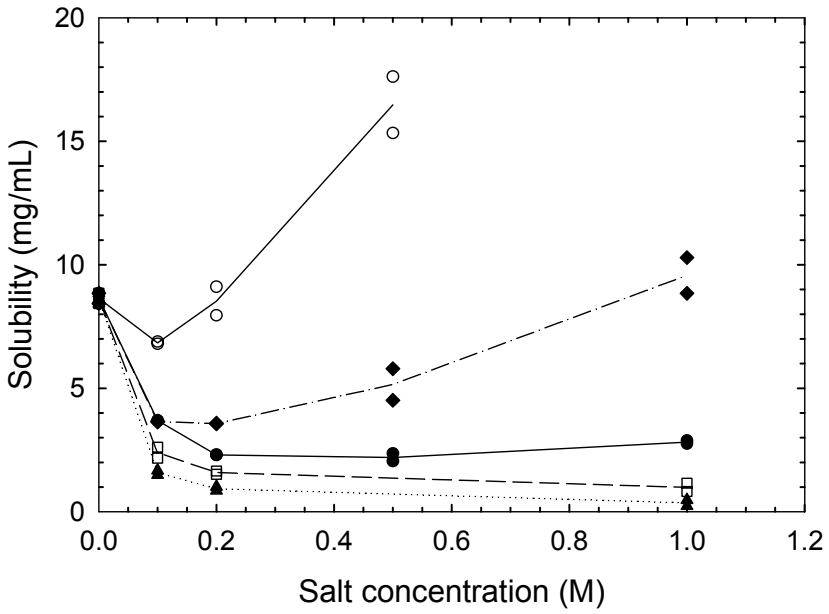

Figure 2.6. Solubility as a function of salt type and concentration. Influence of sodium sulphate (▲, dotted line), sodium acetate (□, dashed line), sodium chloride (●, solid line), sodium nitrate* (◆, dash-dotted line) and sodium thiocyanate (○, solid line) on the solubility of α-amylase at pH 9 ± 0.2 and 40°C after 96 h.**

*** At 0.5 and 1 M the relative standard deviations in the final protein concentration between the solution starting from a higher and a lower initial protein concentration were greater than 10%.** When 0.2 M salt was used, the final pH was 8.7 and 8.6 for crystallisations started with a high and low initial supersaturation, respectively. Using 0.5 M salt led to final pH values of 8.5 and 8.4 from solutions started with a high and low supersaturation, respectively. At 1 M thiocyanate no crystals were obtained.**

The results above are partly in agreement with solubility studies carried out on porcine pancreatic α-amylase with a molecular weight (55 kDa) and pI (pI 5.9) comparable to BHA (Veesler et al., 1996). In that work conducted at pH 8, the same ranking as in the Hofmeister series was also found using potassium salts of chloride, sulphate and thiocyanate at concentration ranges similar to those examined here. Veesler et al. (Veesler et al., 1996) observed that the solubility was significantly enhanced by thiocyanate as compared to the two other salts, however there was no minimum observed for any of the salts. Nevertheless, evidence of solubility minima in the presence of increasing concentrations of salt has been found for other negatively charged proteins. A minimum in the second osmotic virial

coefficient at 0.15 M sodium acetate was measured for apoferritin (Petsev & Vekilov, 2000). The second osmotic virial coefficient can be correlated to solubility (George & Wilson, 1994; Haas et al., 1999), thus a minimum in solubility can also be expected where this coefficient reaches a minimum. Furthermore, the solubility of the negatively charged enzyme, xylose isomerase, passed through a minimum at 0.17 M magnesium sulphate (Vuolanto et al., 2003). Solubility minima for positively charged proteins, such as lysozyme, have not been reported for the anions studied here, but have been for some cations such as magnesium (Broide et al., 1996) and caesium, cobalt, manganese, rubidium and ytterbium (Bénas et al., 2002), typically at concentration ranges above 0.6 M.

Effect of cations on solubility

In analogy to the anion series, cations can also be ranked according to their efficiency to precipitate proteins. In contrast to anions, the influence of cations is generally regarded to be of minor degree. However, the ranking order was found to be independent of the pH at which the crystallisations were carried out relative to the pI of the proteins (Riès-Kautt & Ducruix, 1997), although this was never systematically verified. The different rankings of the cations which were introduced above are based on different experimental conditions with respect to pH and temperature and the corresponding studies were carried out on different proteins, salt type and concentration. Thus those results cannot be directly compared to each other or to the system used in this study, but serve to raise important questions on the mechanism of protein solubility modification by cations. To shed light on those mechanisms with respect to negatively charged proteins, further experiments were conducted with *Bacillus halmapalus* α-amylase. The chosen nitrate salts of caesium, lithium, potassium, rubidium and sodium partly combine the two series examined by Green (Green, 1932) and Bénas (Bénas et al., 2002). The choice of experimental conditions was identical to those used above for the anions (i.e. Figures 2.2 to 2.6) and thus allowed for direct comparison of the influence on the solubility of α-amylase. Monovalent cations were chosen to ensure consistent stoichiometric relations between cations and their nitrate-counterions. Divalent cations are known to exhibit very specific effects on the crystallisation behaviour of e.g., insulin in the presence of Zn^{2+} (Kadima et al., 1993) and lysozyme in the presence of Mg^{2+} (Broide et al., 1996). Moreover, three Ca^{2+}-ions are a fundamental part of the structure of bacterial α-amylases such as the *B.*

halmapalus α-amylase used in this study (Machius et al., 1998). Thus the influence of divalent cations on solubility properties was considered to be too complex to allow for reasonable comparison and was not included in these studies.

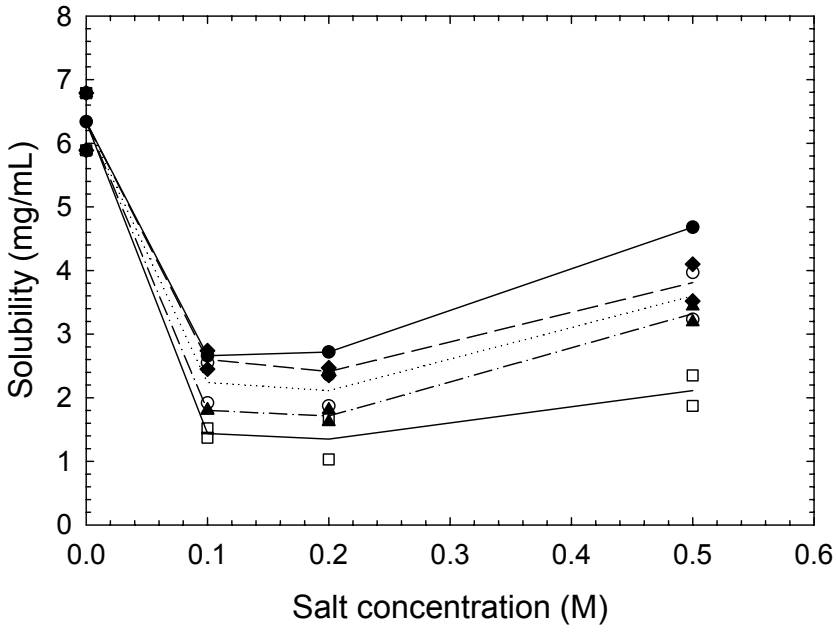

Figure 2.7 Solubility determined by crystallisation as a function of salt type and concentration. Influence of lithium nitrate (□, solid line), caesium nitrate (▲, dash-dotted line), rubidium nitrate (○, dotted line), potassium nitrate (◆, dashed line) and sodium nitrate (●, solid line) on the solubility of □-amylase at pH 9 ± 0.2 and 40°C after 96 hours. Each individual data point represents the average of a double determination.

The five monovalent nitrate salts used showed the same general trends within the concentration range examined (Figure 7) and in all cases, the cations could be ranked from the most efficient precipitant to the least as: $Li^+ > Cs^+ > Rb^+ > K^+ > Na^+$. The solubility decreased by the addition of 0.1 M of each salt. Further increase in salt concentration up to 0.2 M did not lead to significant changes whereas the solubility increased with increasing salt concentration up to 0.5 M. For the cations tested, the relative difference between the least and greatest effect on solubility at a given salt concentration was not more than 2.5 fold. In comparison, anions at the same concentration range as used in Figure 7 induced up to 15-fold differences in the solubility of this enzyme (Figure 2.7). Since anions are capable of altering

the solubility within a significantly broader range than cations, their use for steering solubility would be more appropriate.

The ranking of the cations with respect to their effect on solubility is reversed compared to what was reported by Riès-Kautt (Riès-Kautt & Ducruix, 1997). The only exception is the placement of lithium (Figure 2.7). Neglecting the latter, the results indicate that the Hofmeister series for cations does indeed invert when the net charge of the protein changes. Reasons for the placement of lithium were unclear; however, it was initially suspected that the method of salt addition (as solid) may be partially responsible since insoluble aggregates were only observed upon lithium addition but not for the other salts. Furthermore, the degree of aggregation seemed to depend on how fast lithium nitrate was added. It was thus supposed that the solid salt addition created large concentration gradients during dissolution, which may have been responsible for the aggregate formation. As a consequence, the experiments were repeated by adding appropriate amounts of a 3 M aqueous lithium nitrate stock to the protein solution whilst efficiently mixing. Within the concentration range examined, no visible aggregates formed. However, in an additional experiment, the protein solution was adjusted to contain a concentration of 1 M lithium nitrate. Here, the formation of visible aggregates was observed approximately two hours after salt addition. Nevertheless, the protein solubility data obtained did not show any significant differences between solid and liquid salt addition. The method of salt addition did not affect the enzyme activity, indicating that the protein remaining in solution was not denatured by adding lithium.

As mentioned above, aggregate formation was not only observed following solid salt addition but also after adjusting the lithium concentration of the amylase solution to 1 M with dissolved lithium nitrate. It was thus supposed that smaller and *a priori* invisible aggregates might form in experiments with other lithium concentrations and may help to account for the unexpected ranking of lithium. To further elaborate on this hypothesis, the polydispersity of the supernatant stemming from crystallisation experiments conducted at 0.2 M salt concentration was analysed by dynamic light scattering (Figure 2.8). The polydispersity takes into account how many scattering particles of diameters different from that of the monomeric amylase molecule are present in the solution of interest. Two distinct populations could be identified, one having a mean diameter of approx. 6 nm which corresponds to the size of the protein monomer, and the other with mean diameters between 200 and 1000 nm depending on the salt used. The peaks with size of ~1 nm were considered to be artefacts of unknown

origin. The lowest amount of monomeric form relative to the aggregates was observed with sodium nitrate, whereas in the presence of lithium nitrate the protein was found mainly as the monomer (Figure 8). The type of cation also affected the size of the aggregates formed. Sodium resulted in the smallest aggregates (average size ~200 nm) and lithium and caesium the largest (average ~1000 nm). In fact, the order of the cations effects on aggregate size (Figure 2.8) is the same as that for solubility (Figure 2.7). It is worth noting that the light scattering measurements were carried out below the protein solubility limit after crystallisation had finished and that despite the presence of soluble aggregates, regular crystals had easily formed. It can therefore be concluded that it is not likely that the aggregates act as crystal nuclei. The aggregates observed here appear to be in equilibrium with the monomers and to be formed rapidly, since the solution was passed through a 0.2 μm filter immediately before being subjected to light scattering measurements.

It is generally accepted that increasing protein concentrations favour aggregation and concomitantly increase the polydispersity (Dai et al., 2004; Georgalis & Saenger, 1999). However, the opposite has been observed for porcine pancreatic α-amylase where increased protein concentrations lead to more monodisperse distributions (Veesler et al., 1993). Veesler et al. (Veesler et al., 1993) also found that increasing salt concentrations reduce the polydispersity which is in agreement with data for lysozyme (Dai et al., 2004). On the other hand, insulin formed oligomers at increased salt concentration (Kadima et al., 1993) possibly caused by partial screening of charge repulsion. It thus seems likely that the effects of changing chemical conditions on the polydispersity are protein specific.

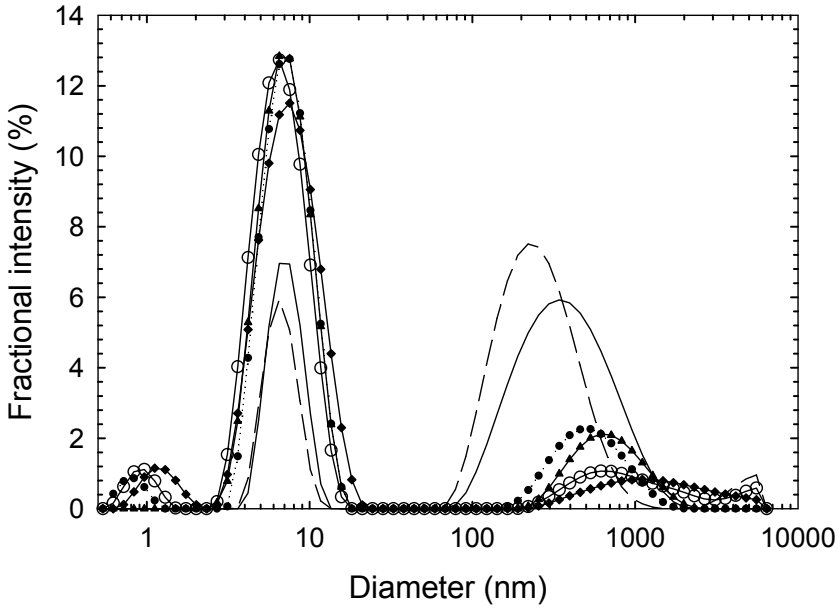

Figure 2.8 Particle size distribution in the supernatant harvested 96 hours after crystallisation of a-amylase was started using 0.2 M of different salts at pH 9 and 40°C. The data is expressed in terms of fractional intensities of the scattered light. No salt (solid black line), sodium nitrate (dashed line), potassium nitrate (●, dotted line), rubidium nitrate (▲, solid line), caesium nitrate (◆, solid line), lithium nitrate (○, solid line), 0.2 M each.

The results described above indicate that, although small in number, the largest aggregates develop in the presence of lithium and caesium salts as compared to other salts during crystallisation, but this finding cannot explain the unexpected positioning of lithium within the Hofmeister series for cations. However the results suggest that the different salts examined may have different effects on the surface properties of the protein. The studies were extended to examine the zeta potential of α-amylase in solutions containing none and 0.2 M of the nitrate salts of lithium, potassium and sodium within a pH-range between 6 and 9. Furthermore, the theoretical charge curve was calculated based on the amino acid sequence and used for comparison. The results (Figure 2.9) show that the zeta potential in solution without added salt changed by ~1 mV in the pH range between pH 7 and 9 whereas it increased dramatically towards the point of zero charge when the pH was lowered from 7 to 6. Similar trends were seen for the calculated charge curve for the enzyme solution free of salt. At all pH values, the magnitude of the zeta potential of the protein was much lower in the presence of lithium than for the other salts, and for the case without salt addition. Thus the

point of zero zeta potential for lithium (pH 6.3) was higher than that predicted for the other salts and salt free solution (~5.8). The latter value corresponds well with the pI of the protein (5.8 - 6.1) determined by isoelectric focusing (Figure 2.1 B) and by the calculated charge curve based on the amino acid sequence (Figure 2.9). The results suggest that the apparent pI was not changed by potassium or sodium nitrate but was by lithium nitrate which may thus account for the lower solubility resulting from this salt compared to the others tested (Figure 2.9).

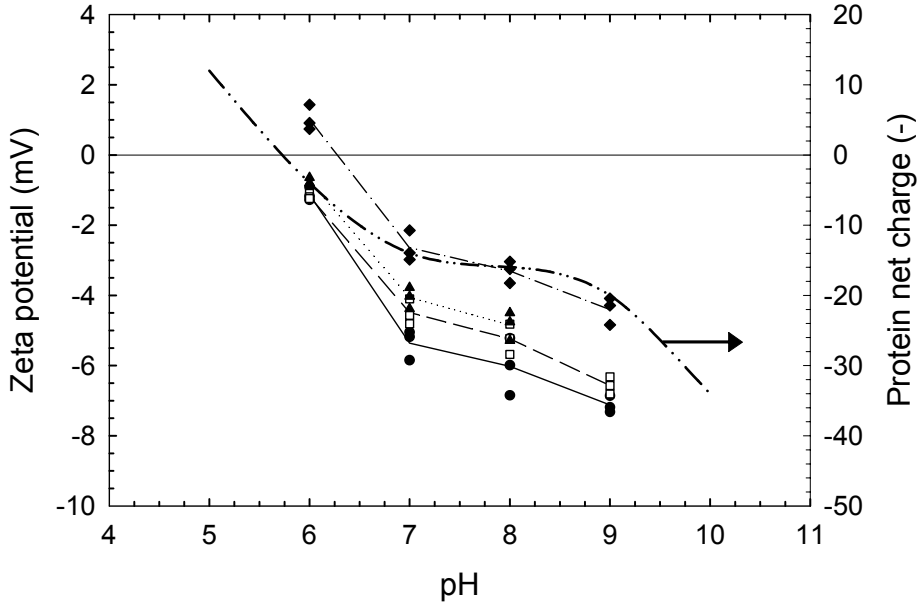

Figure 2.9 Zeta potential and protein net charge of α-amylase as a function of pH in the presence of none and 0.2 M of different nitrate salts. Zeta potential no salt (●, solid line), sodium nitrate (□, dashed line), potassium nitrate (▲, dotted line), lithium nitrate (◆, dash-dotted line). Calculated protein net charge no salt (dash-dot-dotted line).

A number of other workers have documented the effect of salt addition on the apparent pI of other proteins, but no one has demonstrated salt specific effects. In particular the influence of different monovalent cations on the pI of a dissolved protein does not appear to have been the subject of any previous studies. A decrease in the magnitude of the zeta potential (i.e. to lower positive values) caused by sodium chloride, sodium nitrate and sodium thiocyanate was found for lysozyme crystals suspended in 50 mM acetate buffer in the pH range between 3.75 and 5.5 (Lee et al., 2001). In this pH-range, the net charge of lysozyme also decreased within a

comparable degree, emphasising the link between charge curve and zeta potential. The magnitude of the zeta potential also decreased (to less positive values) for a 25 mg/mL lysozyme solution at pH 4.5 buffered in 0.1 M sodium acetate when the sodium chloride concentration was increased from 0 to 0.5 M (Dai et al., 2004). A significant influence of chloride binding on the net charge and the position of the isoelectric point has been demonstrated for ovalbumin and BSA (Windsor, 2003). Furthermore, shifts in pI caused by specific ion adsorption have been observed for alumina based ceramic membranes for which some anions shift the pI towards lower pH values and cations towards higher pH-values (Mullet et al., 1997). Although ion adsorption and ion binding are not the same processes, their influence on the zeta potential of the system is expected to be comparable.

Although no direct links between the net charge or zeta potential and the solubility have yet been established, it is commonly accepted that protein solubility increases with increasing net charge (Riès-Kautt & Ducruix, 1997). Given the strong impact of pH on BHA solubility it thus seems that the reasons for the placement of lithium in the cation series for BHA are due to the ability of lithium to shift the pI to higher values (~0.5 pH-units) and to reduce the magnitude of the zeta potential compared to the other salts studied. The solubility studies in the present work using the cations including lithium were conducted at a pH of 9 and furthermore it was shown that solubility was greatly affected by reducing pH from 10.5 to 8 or 9 (Figure 2.4). Since the calculated charge curve and the zeta potential (Figure 2.9) follow similar trends, it is argued that lithium's ability to lower the magnitude of the zeta potential is also true for the protein net charge. It would thus be expected that the magnitude of the protein net charge at pH 9 would be reduced from -20 to -17 by lithium, which corresponds to a decrease of 15%. All of the other cations tested here had a minor effect on solubility compared to anions (Figure 2.7), thus modification of the protein net charge by 15% would be expected to have a dramatic effect not only on the solubility, but also the position of lithium within the Hofmeister series. We postulate that the reduction in solubility upon lithium addition at pH 9 is not only the result of charge screening but also caused by the capability of the ion to change the pI of the protein by binding to it. In contrast, the other salts tested here did not dramatically affect pI and thus we believe they only modify solubility of BHA by charge shielding. Further work should be conducted to test this hypothesis, for example by studying BHA crystal structure in the presence of lithium, and other salts.

The question arises as to why this effect was not observed for other proteins, e.g., lysozyme. In the case of lysozyme, almost all solubility studies were carried out at pH-values of between 4 and 5, far below the pI of 11.2. It is expected that the effect of changes in pI of e.g., 0.5 pH units may become so small when solubility studies are carried out five or more pH-units below the pI that the ranking of monovalent cations would not be affected. Furthermore, the net charge of lysozyme is much less sensitive towards pH-changes throughout a large pH range, from 4 to 10 and passes through a plateau between pH 6 to 9 (Cacioppo & Pusey, 1991a; Riès-Kautt & Ducruix, 1997) which is quite the opposite for α-amylase (Figure 2.9). Moreover, the solubility of lysozyme at the pH-ranges typically used, i.e. 4 to 5, is also rather insensitive to shifts in pH, particularly at sodium chloride concentrations greater than 0.25 M (Riès-Kautt & Ducruix, 1997). The high pI of lysozyme results in the fact that all of the solubility measurements were conducted at positive protein net charge at which the binding of cations to the molecule is aggravated by electrostatic repulsion. The solubility characteristics of lysozyme may therefore suppress the effects of lithium observed on α-amylase solubility under the conditions studied here. Moreover, the degree by which lithium can shift the pI and the resulting impact on net charge and solubility may be protein specific and could be considerable for some proteins but insignificant for others.

It is often assumed that reduction of net charge on a protein is the usual mechanism underlying the effects of salts on its solubility. In particular electrostatic screening of charged groups, and the adsorption or binding of counter ions by the protein are believed to be the main mechanisms behind the solubility dependence on salt concentration (Retailleau et al., 1997). This concept allows for an explanation of the absence of salting-in effects (i.e. increase in solubility) at low ionic strengths for charged proteins, which is in contrast to the findings of Green (Green, 1932), who observed first an increase followed by a decrease in solubility of carboxyhemoglobin near the pI as the salt concentration was raised. The absence of a salting-in effect at low salt concentrations in lysozyme solubility has been reported for many anions (Retailleau et al., 2002) and would be in good agreement with the findings presented here. It has, however, to be taken into account that the effect of screening and/or adsorption of anions on charged residues of basic lysozyme is not directly comparable to the results obtained for the acidic protein α-amylase. The adsorption of anions would lead to an increase in net charge at pH > pI and a concomitant increase in solubility, which would be in contrast to our findings. It is therefore speculated that at low salt concentrations the dominating effect is

screening of negative patches on the protein surface by sodium ions, which would reduce the charge repulsion between the protein molecules, and thus lead to the observed decrease in solubility. At high salt concentrations the α-amylase's solubility increased again in the presence of chloride, nitrate and thiocyanate, the extent of which differed significantly (Figure 6). This increase might be attributed to the association of these anions with polar or positively charged side-chains, which would result in a rise in net charge, and thus increased repulsion. The degree of this association with anions could be responsible for the specific effect of each given anion on the α-amylase's solubility. It is known that thiocyanate binds 25 times more tightly to human serum albumin, and in greater numbers than chloride ions (Scatchard et al., 1950). A similar effect is likely to occur for BHA, which might explain the strong effect of increasing the enzyme's solubility. Another proposed mechanism affecting the solubility is the accumulation of hydrated sodium ions in the vicinity of the protein surface at higher salt concentrations. The hydration forces so induced, would also be expected to result in increased repulsion between the protein molecules, thereby raising their solubility (Petsev & Vekilov, 2000). Petsev & and Velikov (Petsev & Vekilov, 2000) point out that the origin of hydration forces is still not clearly identified, and that the interaction between hydration and electrostatic forces appears to be system specific. For BHA studied here, the effect of monovalent cations was found to be of minor importance in comparison to anions within a concentration range from 0 to 0.5 M. The results suggest a dominating role of anions whereas cations only induce subtle effects on the solubility (Figure 2.6 and Figure 2.7). This finding would emphasise that electrostatic screening dominates until 0.1 to 0.2 M where the solubility decreases. At higher concentrations repulsive forces caused by accumulation of counter-ions cause an increase in solubility. In light of this, the reversal of the Hofmeister series found for cations seems reasonable.

2.5 Conclusions

The solubility behaviour of *Bacillus halmapalus* α-amylase has been characterised, and pH changes, salt type and concentration exert the greatest impact on the solubility of this enzyme. Only a minor retrograde temperature dependence of the solubility over the range of 5°C to 40°C occurs. The lowest solubility (0.1 mg/mL) is at pH 6.1 (close to the isoelectric point of ~pH 5.8) and in the presence of 1 M sodium sulphate at pH 9 at a temperature of 40°C. The highest solubility was found at 0.5 M sodium thiocyanate (16 mg/mL), also at pH 9 and 40°C.

The Hofmeister series was confirmed for anions and, with the exception of lithium, also for cations. Lithium resulted in the lowest solubility of BHA for the cations examined, which was concluded to be due to binding of the cation to the protein molecule leading to a shift in pI from 5.8 to 6.3. This was proposed to reduce the net charge of the protein more than expected by charge shielding, which leads to reduced solubility due to the strong impact of pH changes on solubility. Further examinations revealed that compared to other monovalent nitrate salts, the polydispersity was more reduced in BHA solutions containing lithium nitrate. The shift in pI along with the reduced polydispersity of BHA could be the reason for the unexpected position of lithium within the Hofmeister series for cations.

3 Development of crystal size distributions in batch crystallisation processes of a recombinant *Bacillus halmapalus* α-amylase

Cornelius Faber,[a,b] Timothy J. Hobley,[b] Jørgen Mollerup,[c] Svend G. Kaasgaard,[a] and Fabienne Espitalier [d,*]

[a] Novozymes A/S, Novo Alle, DK-2880 Bagsværd, Denmark

[b] Center for Microbial Biotechnology, BioCentrum-DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[c] Department of Chemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[d] Laboratoire de Génie des Procédés des Solides Divisés, Ecole des Mines d'Albi-Carmaux, Allée des Sciences, 81013 Albi Cedex 09, France

*Corresponding author

e-mail: espitali@enstimac.fr (F. Espitalier)

Telephone: +33 563 49 31 51

Fax: +33 563 49 30 25

Synopsis: The temporal development of crystal size distributions in a batch crystallisation process of *Bacillus halmapalus* α-amylase is studied and discussed.

Keywords: amylase; crystallisation; crystal size distributions; electrical sensing zone method, proteins; supersaturation;

3.1 Abstract

The temporal development of the supernatant protein concentration, the crystal size, size distribution and number of batch crystallisation processes of a recombinant *Bacillus halmapalus* α-amylase conducted at 40°C under constant mixing in 1.5 mL-Eppendorf tubes was followed. For all the three different initial supersaturation ratios tested, the size distribution did not notably change and the crystals grew very rapidly to a mean number diameter of between 4 and 5 µm. The growth was too fast to be resolved by the applied measuring technique. Independent of the three different initial supersaturation ratios tested, approximately 10^7 crystals/kg were found immediately after induction of the supersaturation. Within the following two to three hours, the crystal concentration increased from 10^7 to 10^9 crystals/kg without significantly reducing the protein concentration in the supernatant. Depending on the supersaturation, the crystal concentration stayed constant for a certain time. When an initial supersaturation ratio of 6.6 was used, the crystal concentration started increasing after four hours and remained constant between 24 and 48 hours and $1.2x10^{11}$ crystals/kg were formed during the process. The crystal concentration of the process started from an initial supersaturation ratio of 5.5 was notably increased after six hours and was still ongoing after 29 hours and at this time, $1.4x10^{10}$ crystals were counted. An initial supersaturation ratio of 3.8 was not high enough to initiate a crystallisation process within 22 hours but $7.8x10^8$ crystals/kg were present throughout the entire process.

3.2 Introduction

The emerging interest in bulk protein crystallisation in product recovery and formulation in core industrial areas, such as the enzyme and pharmaceutical industry, has increased the need for a better process control to efficiently produce crystals of desired properties (Margolin, 1996). Data on protein crystal growth kinetics and phase diagrams under process-relevant conditions are sparse since the majority of studies of protein crystallisation are driven by the need to produce high quality crystals suitable for structural protein studies by X-ray diffraction (Judge et al., 1995). The results of crystal growth kinetics derived from studies

conducted by growth techniques typically employed such as vapour diffusion in hanging or sitting drops (Forsythe & Pusey, 1994) have only little relevance in bulk protein crystallisation which are often conducted in ton-scale, typically in batch mode. Whilst crystal perfection is critical for the generation of meaningful diffraction patterns prerequisite for high resolution protein structure determination, large crystals of compact shape and homogeneous in size grown within a short time (ideally less then 24 hours) and at high yields are desirable in downstream processing (Rohani et al., 1990). The process design typically employed in crystal growth studies on inorganic or small organic molecules is to a large extent congruent to the one used in bulk protein crystallisation (Tavare, 1986) but comparison is hampered by big differences in the habit between crystals consisting of inorganic or small organic and protein molecules. The much higher solvent content (up to 80%) and the higher supersaturation (typically one order of magnitude higher) required to grow protein crystals are only two examples of key differences in the crystal habit between inorganic and protein crystals (Chernov, 1997). Another aspect is that batch crystallisations performed in crystallisers similar to the ones used in production require a minimum working volume in the range of 200 to 1000 mL (Jacobsen et al., 1998) which may lead to very high experimental costs due to the large demand of proteins, which may often be available in minute quantities (Schmidt et al., 2005). Neither the traditional crystallisation of small inorganic molecules nor the protein crystallisation employed for X-ray analysis can be expected to be similar to bulk protein crystallisation but they may provide worthwhile guidelines. The crystallisation process of a recombinant *Bacillus halmapalus* α-amylase (BHA) studied was conducted in batch mode but in comparatively small working volumes (i.e. 1 mL). The only parameter to be varied was the supersaturation to study its influence on the temporal development of the crystal size, distribution and number as well as the supernatant protein concentration. BHA was chosen as a model system because its solubility and phase diagrams were a priori available and it readily crystallises. The solubility curves were established by classical batch crystallisation methods identical to the ones studied here. In the solubility studies, it was found that equilibrium had sufficiently been approached after 96 hours but no further attention was given to kinetic aspects of the process, e.g., how quickly the crystals were formed and how many (Chapter 2). This study consequently aims to describe the crystallisation process employed to determine the equilibrium concentrations in quantitative terms. The knowledge generated may provide useful guidelines for the design of large scale crystallisation processes of bulk proteins.

3.3 Materials and methods

3.3.1 Preparation of *Bacillus halmapalus* α-amylase

A recombinant *Bacillus halmapalus* α-amylase was expressed in *Bacillus licheniformis* cultivated on a complex medium, and purified at Novozymes A/S, Bagsværd, Denmark. Cells and other solids were removed by filtration, and the resulting solution was concentrated by ultrafiltration (10 kDa cut off), before crystallising the enzyme by lowering the pH to 7.5. The crystals were collected by centrifugation. For all experiments, a buffer containing 30 mM boric acid of analytical grade (AppliChem, Darmstadt, Germany) was used. The reagent was dissolved in ultra-pure water of 18.2 MΩcm resistivity (Millipore, Bedford, MA, USA). Crystals of α-amylase were then dissolved in the buffer at pH 11 where the BHA solubility is high by adding 1 M sodium hydroxide of analytical grade (Merck, Darmstadt, Germany). After dissolution, the solution was filtered through a 0.22 μm pore size cellulose acetate filter (Sartorius, Göttingen, Germany). Three different stock solutions of approximately 60, 50 and 35 mg/mL protein concentration were prepared. The supersaturation necessary for initiating the crystallisation process was induced by slowly reducing (approximately over 5 minutes) the pH to the value of interest, i.e. pH 9 using 0.1 M acetic acid of analytical grade (Merck, Darmstadt, Germany). The solution was transferred into 1.5 mL safe-lock PCR-clean Eppendorf tubes (order no. 0030 123 328, Eppendorf, Hamburg, Germany) using a sample volume of 1 mL. The tubes were closed and placed on a thermo mixer (Model 5355, Eppendorf, Hamburg, Germany) that permitted precise temperature control (40 ± 1°C) and suitable mixing conditions (1400 rpm, eccentricity of 3 mm). The pH was not adjusted during the crystallisation process since previous experiments showed that the variation was less than ±0.2 units.

3.3.2 Analysis

At each given experimental time, a tube was removed from the mixer and the content filtered via a 0.22 μm cellulose acetate filter (Sartorius, Göttingen, Germany). The filtrate was analysed for its protein concentration by reading the absorbance at a wavelength of 280 nm using an HP 8452A spectrophotometer (Hewlett Packard, Palo Alto, CA, USA). To ensure

comparability with previous solubility studies, the absorbance values were converted into BSA-equivalents (Chapter 2).

The crystallisation process in terms of crystal number and size was followed by electrical sensing zone measurements using a Beckman Coulter Multisizer II (Beckman Coulter, Fullerton, CA, USA). An orifice of 100 µm diameter was used which has a measuring range of between 2 and 67 µm. Instead of the Isoton electrolyte solution as recommended by the manufacturer (Beckman), the crystals were counted and sized in a 30 mM boric acid buffer saturated with BHA at pH 9 to prevent crystal dissolution during measurements (equal to crystallisation conditions). The solution saturated with protein had a conductivity of 0.83 mS/cm whereas the conductivity of the Isoton solution was 11.1 mS/cm which could not be used because of strong protein precipitation. Accurate size distribution measurements on protein crystals using the electrical sensing zone method in a solution different from Isoton of a conductivity of approximately 2.2 mS/cm were reported (Jacobsen, 1998). The saturated protein solution was filtered via a 0.45 µm a PTFE membrane (Millipore Corp., Bedford, MA, USA). Latex beads of 3 µm diameter were used to calibrate the system and a cell constant K_d of 993 was found. Verification measurements of latex particles of 15 µm diameter confirmed high accuracy of the system using the determined cell constant. Before each measurement, approximately 50 mg of saturated protein solution was transferred to a beaker, the net mass weighted out and the background noise determined by triple measurements. After this, the remaining mass of the saturated protein was noted down to account for the mass withdrawn during the background measurements (2 mL sample volume was needed per measurement). The crystal suspension was then transferred into the beaker and the added mass recorded. In the early stages of the crystallisation process the entire content of one Eppendorf tube (1 mL) was added whereas in the later stages of high crystal density as little as 50 to 100 µL was found to be sufficient. The adjustment of the crystal concentration during the process was necessary to minimise interferences caused by primary (when two or more crystals pass through the orifice in close proximity, which may cause overlapping signals interpreted by the device as resulting from a single crystal) and secondary coincidence (when the crystals pass through the orifice almost simultaneously, so that the resulting pulse is the sum of the pulses the crystals would have caused individually). The crystal size and number was determined in triplicates. One set of measurements was finished within 10 minutes after the crystals were transferred into the saturated buffer. Within this time, no crystal dissolution was noted since the size distributions as well as the number of crystals were essentially constant. Per size

interval, the averaged concentration of the background was subtracted from the average crystal concentration.

To cross-check the results obtained from the Beckman Coulter Multisizer II some crystals grown from an initial protein concentration of 60 mg/mL were collected by passing the suspension over a glass frit covered with a 0.45 µm PTFE membrane (Millipore Corp., Bedford, MA, USA) after the crystallisation process had approached equilibrium. The crystals were then briefly dried at room temperature for approximately two hours and analysed by a Philips XL30 ESEM-FEG environmental scanning electron microscope (Philips Electron Optics, Eindhoven, The Netherlands) using a poor vacuum (around 15 mbar).

The results obtained from electrical sensing zone measurements were compared with size distributions of BHA crystals grown under similar conditions (initial protein concentration 70 mg/mL, MES-HEPES-boric acid buffer (10 mM each) and room temperature) which were determined by low angle laser light scattering using a Malvern Mastersizer 2000 (Malvern Instruments, Malvern, UK).

3.4 Results

3.4.1 Crystallisation process started from an initial protein concentration of approximately 61.7 ± 0.5 mg/mL (in the following referred to as CP01)

The initial protein concentration used in this process accounts for an initial supersaturation ratio of 6.6 (defined as the protein concentration in solution c divided by the equilibrium concentration c_s which is of about 9 mg/mL at pH 9) (Feher & Kam, 1985). The first measurement of the crystal concentration which was conducted 10 minutes after generation of the supersaturation indicated the presence of 1.27×10^7 crystals/kg of a size of the measuring range between 2 and 60 µm. It further increased to approximately 10^8 crystals/kg within the next hour before it rapidly raised to 5×10^{11} crystals in the following five hours. While the crystal concentration was increasing the protein concentration in the supernatant was declining (Figure 3.1) which is typical for phase separation processes in which protein molecules get transferred from the liquid into the solid crystalline phase. This phase

separation process was slowed down after seven hours and later on, both the protein concentration in the supernatant as well as the crystal concentration underwent only minor changes and stayed at approximately $2x10^{11}$ crystals/kg. The crystallisation process had almost reached equilibrium after 48 hours as the protein concentration in the supernatant of approximately 9 ± 0.5 mg/mL almost equalled the one measured in solubility experiments under identical conditions after 96 hours (Chapter 2). The mean number diameter d (1,0) almost stayed constant throughout the entire crystallisation process and accounted for approximately 4.2 to 4.6 µm. This finding was also reflected by the crystal size distribution expressed in number of crystals per size interval measured at different times during the process (Figure 3.2). The majority of crystals were always found in the same size interval between 3 and 4 µm (mode number diameter, different to the mean number diameter due to non-symmetrical size distributions).

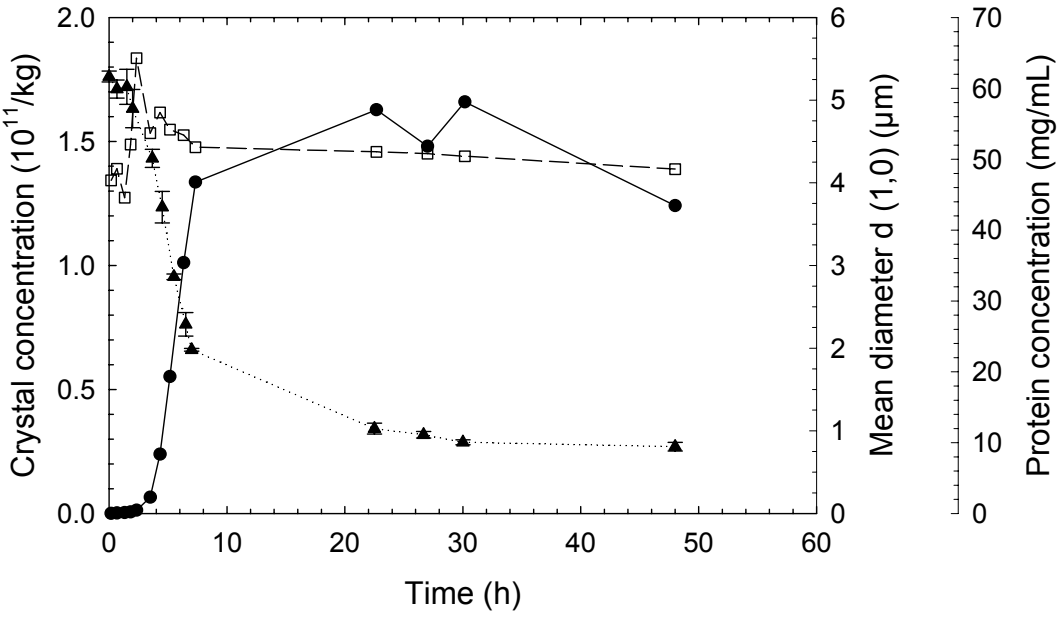

Figure 3.1 Crystallisation process CP01 started from an initial BHA concentration of 60 mg/mL conducted at pH 9 and 40°C as a function of time. (●, solid line): crystal concentration; (□, dashed line): mean diameter d (1.0); (▲, dotted line): protein concentration in the supernatant. The employed initial supersaturation ratio equals a relative supersaturation of 6.6.

The 3-dimensional graph in Figure 3.2 clearly indicates how quickly the crystal concentration was increasing within the first six to seven hours but then entered into a regime of only very

modest changes. Moreover, the graph also shows that the width of the size distribution hardly changed and that there was no tendency towards the growth of larger crystals. Already in very early stages, the size distribution developed its characteristics which were maintained throughout the entire process.

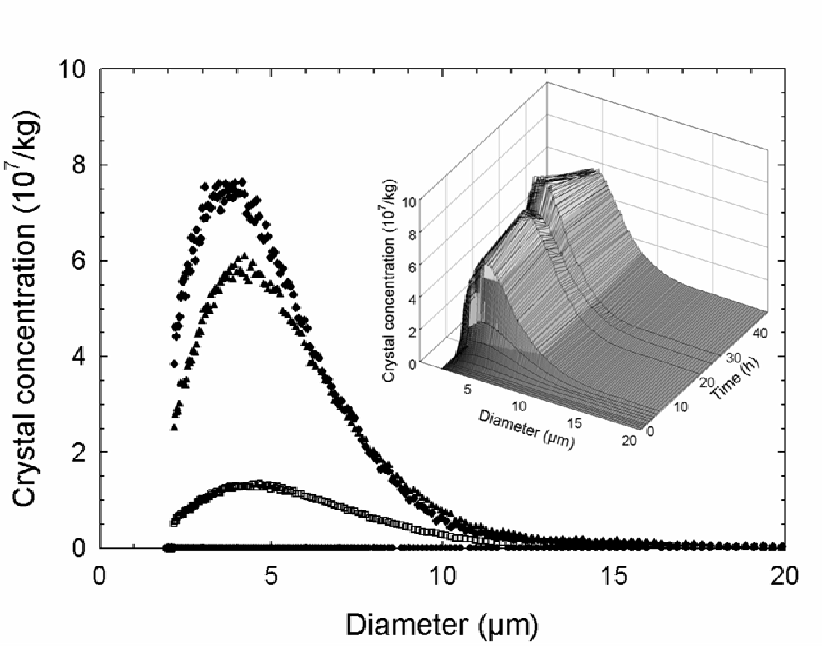

Figure 3.2 Crystal size distribution of CP01 determined by electrical sensing zone measurements at different times after induction of the supersaturation. (●): after 0 hours; (□): after 5.2 hours; (▲): after 23 hours; (◆): after 48 hours grown at 40°C. The initial BHA concentration was 60 mg/mL which equals a relative supersaturation of 6.6. The crystallisation process was conducted at pH 9 buffered in 30 mM boric acid. The 3-dimensional graph represents the complete data set.

To verify the accuracy of the results obtained from the electrical sensing zone measurements, pictures of the crystals were taken by environmental scanning electron microscopy (ESEM). A representative image of crystals grown for 48 hours at pH 9 started from an initial protein concentration of 60 mg/mL is given in Figure 3.3. The picture suggests that only very few crystals were larger than 20 µm but the majority was considerably smaller ranging from approximately 5 to 10 µm. Moreover, the image indicates that only to a minor degree crystal fragmentation was occurring during the crystallisation process as some crystals seem to have broken up into smaller parts whereas other crystals show traces of blistering of smaller crystalline units. Both breakage mechanisms result in fragments of sizes of the same order of magnitude of the parental crystal. Attrition which rounds off the edges of the crystals does not

seem to have been taking place during the crystallisation process. Both fragmentation and attrition take place in strongly agitated systems where the turbulent eddies are smaller than the crystals. In contrast to crystal breakage, attrition is likely to occur in suspensions of very high crystal concentration (Synowiec et al., 1993) which did not seem to have been given in the systems investigated in the current study.

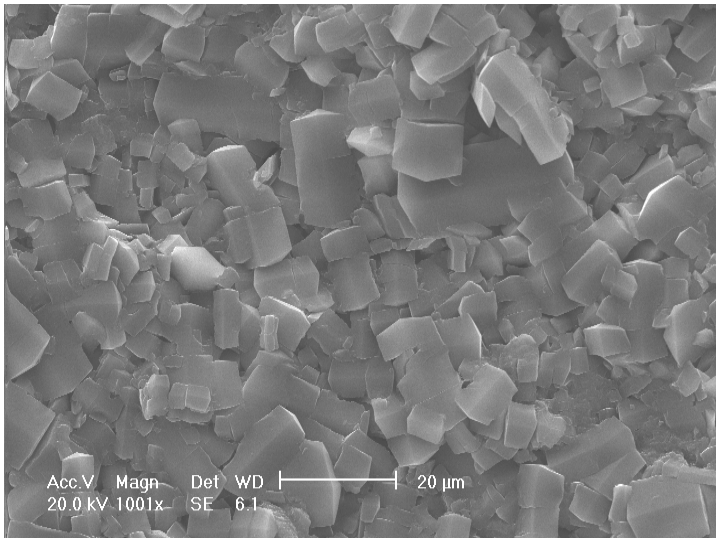

Figure 3.3 BHA crystals grown for 48 hours at pH 9 and 40°C from an initial protein concentration of 60 mg/mL (CP01) seen by environmental scanning electron microscopy (ESEM).

In a separate experiment, the size distribution of BHA-crystals grown for 48 hours at pH 9 and room temperature but from a higher initial supersaturation ratio of 7.7 was analysed by low angle laser light scattering measurements. The resulting size distribution is presented in Figure 3.4 in which the size distribution of CP01 after 48 hours measured by the electrical sensing zone method is given for comparison. Although not quite grown under identical conditions, the data suggests that the two measurements lead to comparable results. The size distribution obtained from low angle laser light scattering is broader and comprises more crystals in the range between 10 and 20 µm whereas more crystals of a size range between 2 and 4 µm were detected by electrical sensing zone measurements. Subsequently, the resulting mean number diameter d 1,0 differs and accounted for 4.7 µm in the electrical sensing zone method whereas a mean number diameter of 6.2 µm was suggested by low angle laser light scattering measurements. Since the crystals were grown at slightly different experimental conditions, i.e. different initial supersaturation ratios and different temperatures, no

conclusive statement can be given whether the varying size distribution are caused by the measuring techniques or by the variations of the crystal growth conditions. The orifice used for the electrical sensing zone measurements had a measuring range of between 2 and 60 µm. Particularly in early stages of the crystallisation process, there is the risk that a considerable amount of crystals remain undetected since they are smaller than the lower limit of the measuring range. Towards the finalisation of the crystallisation process, crystals smaller than 2 µm seem to be present only in small amounts so that the size distributions are correctly presented. However, the comparison of the size distribution obtained from the electrical sensing zone method with the images generated by ESEM and size distributions measured by multi angle laser light scattering suggests that the experimental settings chosen for the electrical sensing zone method leads to results of adequate accuracy.

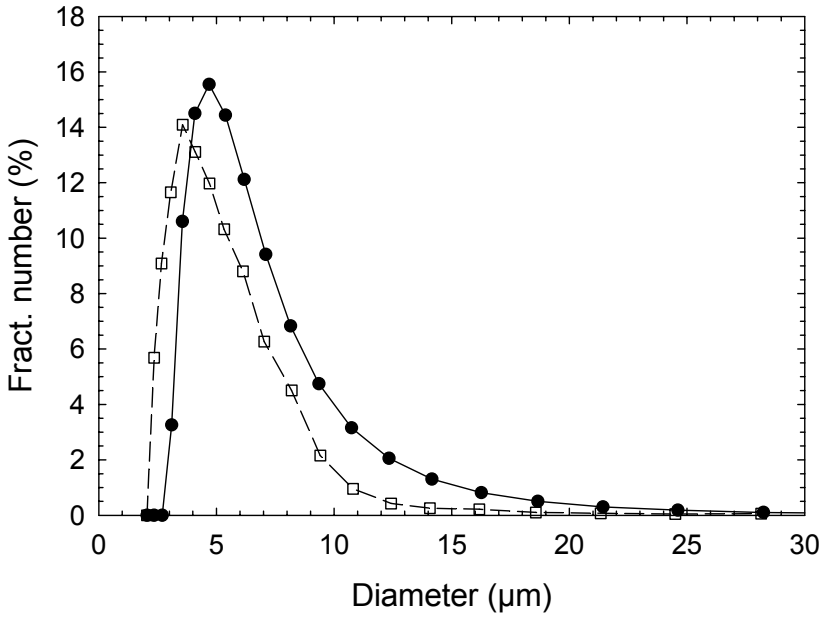

Figure 3.4 Comparison of BHA-crystals grown at pH 9 measured by low angle laser light scattering (●, solid line: Malvern Mastersizer 2000, grown at room temperature) and electrical sensing zone method (□, dashed line: Beckman Multisizer, grown at 40°C) given in fractional number. The mean diameter obtained from low angle light scattering measurements was calculated to be 6.2 µm whereas a mean number diameter of 4.7 µm was found by electrical sensing zone measurements.

Drawbacks of the low angle laser light scattering measurements are that they only provide fractional and not absolute numbers so that no information is given how many crystals are forming during the process. Moreover, this method faces problems once the obscuration caused by increasing crystal densities becomes too high, which leads to inaccuracies in the resulting size distributions. Thus, the electrical sensing zone method was taken as the more appropriate method for the system under investigation and exclusively used in the current study.

3.4.2 Crystallisation process started with an initial protein concentration of approximately 50.4 ± 0.9 mg/mL (in the following referred to as CP02)

A BHA batch crystallisation process started from an initial protein concentration of 50.4 ± 0.9 mg/mL is characterised in Figure 3.5. The initial protein concentration used in this crystallisation process corresponded to an initial supersaturation ratio of 5.5. The basic trends were the same as discovered for CP01 presented in Figure 3.1. After 10 minutes, $7x10^6$ crystals/kg of sizes between 2 and 60 μm were found. Within the next two to three hours, the crystal concentration increased to approximately $2x10^9$ crystals/kg and remained constant for three more hours. Only after six hours, the number of crystals was substantially increasing accompanied by a decrease in protein concentration in the supernatant. After 29 hours the crystallisation process was not yet finished since the protein concentration was still decreasing and the crystal concentration increasing. The protein concentration determined after 29 hours was approximately 13 mg/mL whereas the equilibrium concentration would be expected to be approximately 9 ± 0.5 mg/mL (Chapter 2). After 29 hours $1.4x10^{10}$ crystals/kg were found which is about one order of magnitude less than in CP01. The mean number diameter was found to be only slightly higher than in the previous crystallisation process (approximately between 4.4 and 4.8 μm compared to between 4.0 and 4.4 μm for the previous crystallisation). The mode number diameter was also higher since the majority of particles were found in a size interval between 4 and 5 μm (between 3 and 4 μm in CP01) (Figure 3.2 and Figure 3.6). The size distribution was as broad as in CP01 and did not significantly change during the experimental run time (due to the different scaling of the y-axis the amount of particles on the right end of the distribution appear over-represented (Figure 3.2 and Figure 3.6)).

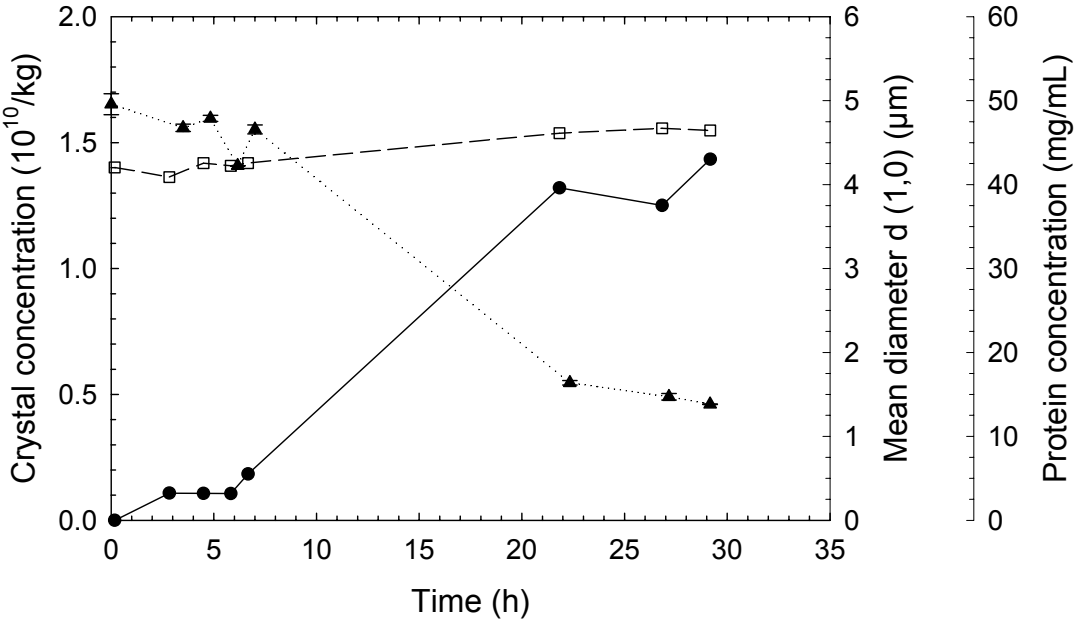

Figure 3.5 **Crystallisation process CP02 started from an initial BHA concentration of approximately 50 mg/mL conducted at pH 9 and 40°C as a function of time. (●, solid line): crystal concentration; (□, dashed line): mean diameter d (1,0); (▲, dotted line): protein concentration in the supernatant. The employed initial supersaturation ratio equals a relative supersaturation of 5.5.**

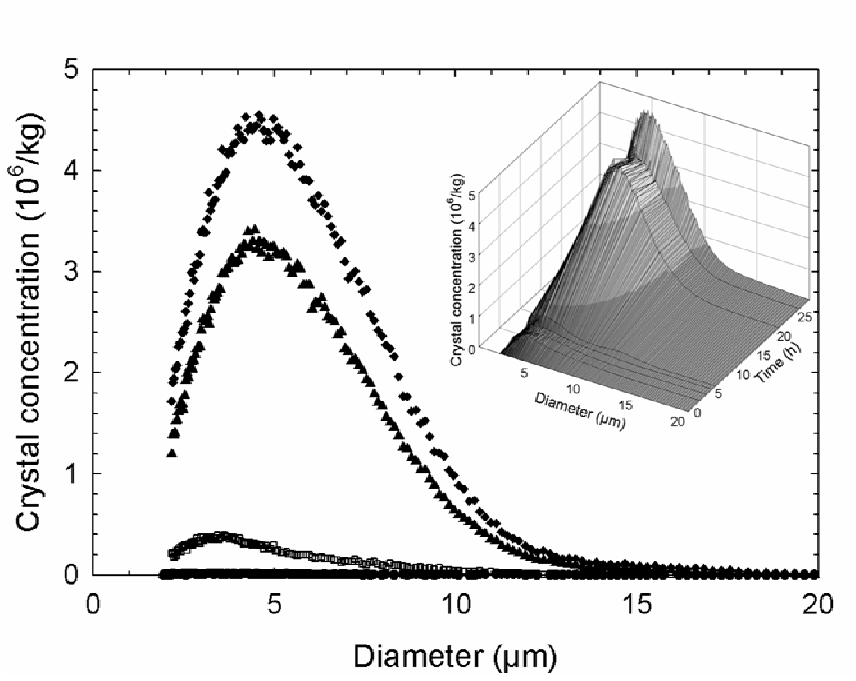

Figure 3.6 **Crystal size distribution of CP02 determined by electrical sensing zone measurements at different times after induction of the supersaturation. (●): after 0 hours; (□): after 2.8 hours; (▲): after 21.8 hours; (◆): after 29.2 hours. The initial BHA concentration was 50 mg/mL which equals an initial supersaturation ratio of 5.5. The 3-dimensional graph represents the complete data set.**

3.4.3 Crystallisation process started with an initial protein concentration of approximately 35 mg/mL (CP03)

The initial protein concentration used in this crystallisation process corresponded to a relative initial supersaturation ratio of 3.8. Similar to CP01 and 02, $6x10^6$ crystals/kg were found after 10 minutes. Within the next two to three hours, the crystal concentration increased to $5x10^8$ crystals/kg (Figure 3.7). In the remaining experimental run time of around 24 hours, both the protein and the crystal concentration almost stayed constant. The maximum crystal concentration of $7x10^8$ crystals/kg was two orders of magnitude less than encountered in CP02 and three orders of magnitude less than CP01. The decrease of the protein concentration between 22 and 27 hours and the concurrent increase in crystal concentration may indicate the onset of the crystallisation process. The crystals had a mean number diameter of 4 µm which was slightly less than determined for the previous crystallisation processes (Figure 3.2 and Figure 3.6). The mode diameter was found to be of between 3 and 4 µm (Figure 3.8) and the size distribution was slightly narrower than observed for the crystallisation processes 01 and 02 and did not change throughout the experimental time.

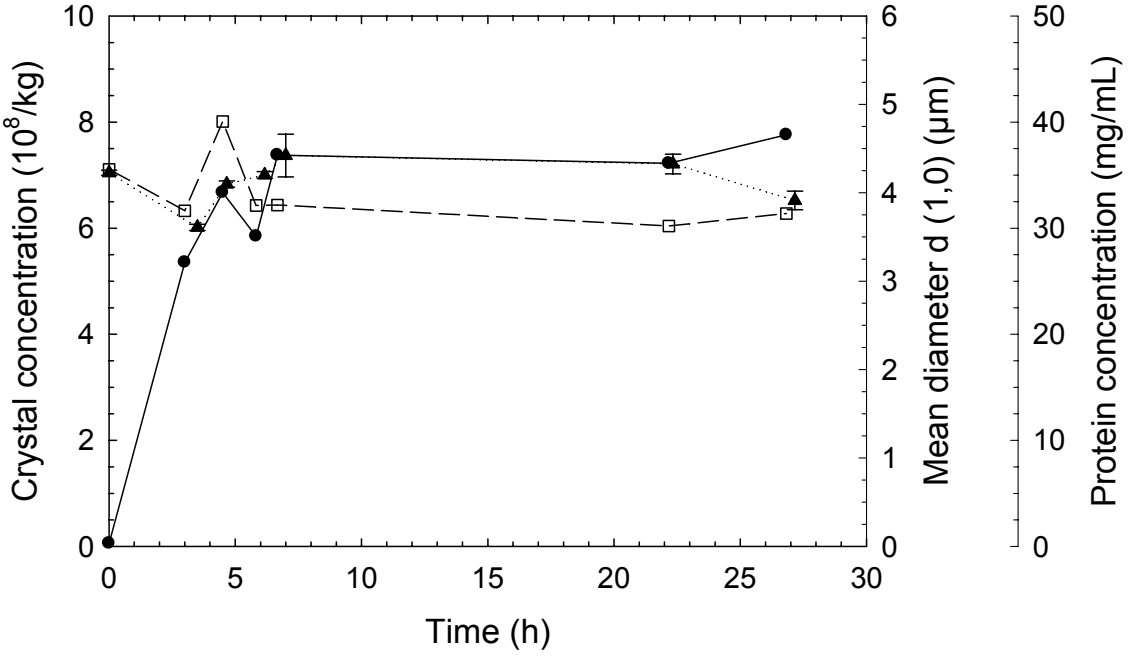

Figure 3.7 Crystallisation process CP03 started from an initial BHA concentration of approximately 35 mg/mL conducted at pH 9 and 40°C as a function of time. (●, solid line): crystal concentration; (□, dashed line): mean diameter d (1,0); (▲, dotted line): protein concentration in the supernatant. The employed initial supersaturation ratio equals a relative supersaturation of 3.8.

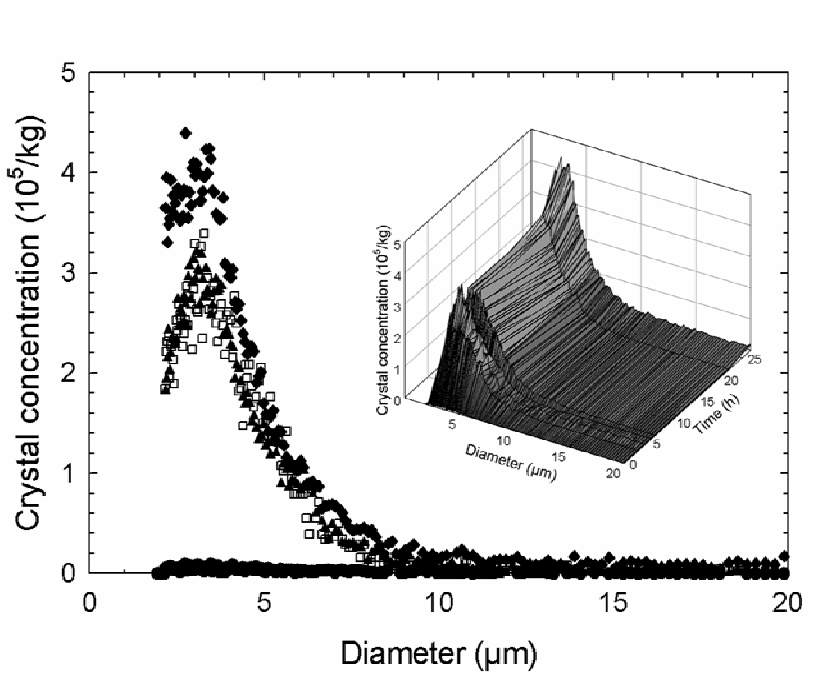

Figure 3.8 Crystal size distribution of CP03 determined by electrical sensing zone measurements at different times after induction of the supersaturation. (●): after 0 hours; (□): after 4.5 hours; (▲): after 22.2 hours; (◆): after 26.8 hours. The initial BHA concentration was 35 mg/mL which equals an initial supersaturation ratio of 3.8. The 3-dimensional graph represents the complete data set.

Table 3.1 Crystal and supernatant protein concentrations at different times of the three examined crystallisation processes

CP01			CP02			CP03		
time (h)	crystal concentration (number/kg)	protein concentr. (mg/mL)	time (h)	crystal concentration (number/kg)	protein concentr. (mg/mL)	time (h)	crystal concentration (number/kg)	protein concentr. (mg/mL)
0.17	1.27E+07	62.2	0.17	7.46E+06	48.7	0.17	5.62E+06	35.05
0.67	2.04E+08	58.99	2.83	1.08E+09	46.45	3	5.35E+08	30.28
1.33	3.64E+08	58.46	4.5	1.07E+09	48.14	4.5	6.67E+08	34.36
1.83	5.83E+08	55.24	5.83	1.06E+09	42.24	5.83	5.84E+08	34.78
2.33	1.36E+09	49.22	6.67	1.84E+09	46.2	6.67	7.38E+08	35.42
3.5	6.55E+09	44.8	21.8	1.32E+10	16.2	22.17	7.22E+08	36.7
4.33	2.39E+10	33.68	26.8	1.25E+10	14.99	26.83	7.76E+08	31.99
5.17	5.52E+10	25.52	29.1	1.43E+10	13.84			
6.33	1.01E+11	22.96						
7.33	1.34E+11	11.44						
22.6	1.63E+11	11.43						
27	1.48E+11	10.3						
30.1	1.66E+11	9.8						
48	1.24E+11	9.01						

3.4.4 Modelling

The concentration profile and the crystal concentrations of CP01 and CP02 were incorporated into population and mass balances in order to predict the resulting crystal diameters. The results are presented in Appendix I.

3.5 Discussion

Probably the most interesting fact was that the measurements suggested that the mean number diameter and the size distributions of the crystals formed did not change during the crystallisation processes examined. Only the crystal concentration increased until the supersaturation was decomposed and the equilibrium concentration reached. Within the tested initial supersaturation ratios the mean number diameter of the crystals ranged between 4 and 5 μm which is small compared to the crystals of a microbial lipase which had a mean volume diameter of 16 μm but of much broader size distribution (Jacobsen et al., 1998) or ovalbumin crystals of 40 μm mean number diameter (Judge et al., 1995), both crystallised in batch mode

and in litre-scale. The crystal growth seems to be a rapid process since no steady increase of the mean diameter during experimental time could be resolved. The growth cessation at small crystal sizes was probably caused by the adsorption of impurities on the crystal surface or the accumulation of lattice errors which prevented further incorporation of protein molecules into the crystal lattice (Weber, 1991). Another reason for the small mean size of the BHA-crystals could have been crystal breakage due to the high mixing frequency and the resulting crashes either between two crystals or between crystal and container wall (Synowiec et al., 1993). Since the BHA-crystals are not spherical (Figure 3.3), they become more fragile with increasing size and are more likely to break. The probability of crystal breakage also gets higher with increasing crystal number. The image in Figure 3.3 which is taken at the end of a CP01 which resulted in the highest crystal concentration of all processes examined, suggests that crystal breakage was occurring only to a minor extent but attrition can be excluded which would have been indicated by rounded-off edges of the crystals (Synowiec et al., 1993). In early stages of the crystallisation process, crystal breakage can probably be neglected such that the crystal size is determined by growth cessation as discussed above. As intact crystals having ceased to grow cannot incorporate further protein molecules, they can be regarded as inert particles since they do not influence the crystallisation process anymore. As a consequence, new crystals must form as the only way to reduce the supersaturation leading to an increase in crystal concentration, either by primary or secondary nucleation, the latter, however, probably playing a minor role.

Unexpectedly, a concentration of approximately 10^7 crystals/kg was found in the first measurement in each of the crystallisation processes examined 10 minutes after the induction of the supersaturation (Table 3.1). This number seems to have been independent of the supersaturation employed. Since all results were corrected for the background and the mean diameter and size distribution was essentially the same as the one determined at later stages of the process, it can be concluded that these particles were crystals. Relative to the amount of crystals counted, dust or other impurities were probably present only in negligible concentrations. It is suggested that the crystals were formed during pH-adjustment. The addition of acetic acid, although of relatively low concentration (0.1 M), may have created pH-gradients while the drop was intruding into the protein solution. Subsequently, an area of high supersaturation was generated until the acid was well distributed throughout the solution. In this area, crystals may have formed which were too low in number to create any turbidity or the typical shampoo look so that the solution appeared crystal-free, which was obviously

not the case. Crystals once formed could not dissolve since the solution was supersaturated and thus must have constituted for the amount of particles counted. As the crystals had reached the final mean diameter, they are assumed as inert particles not further interfering with the crystallisation process for reasons discussed above. An increase in crystal concentration from 10^7 to 10^{8-9} crystals/kg within the first three hours of the crystallisation process was noted, before it either almost immediately further increased (CP01), or remained constant for a certain time before further increasing (CP02) or remained constant during the entire observation time (CP03). The local supersaturations generated during pH-adjustment only lasted for a very short time such that the increase from 10^7 to 10^{8-9} crystals/kg could not be a result of these supersaturations anymore. This increase could be an artefact caused by the measuring principle. As the detection range was between 2 and 60 µm, a certain amount of crystals formed in early stages may have been smaller than 2 µm and have thus remained undetected. Only after a certain time, they reach the detectable size and cause the increase in crystal concentration. The nucleation rate is typically more dependent on supersaturation than the growth rate. For lipase, the power dependence on supersaturation of the nucleation rate was 10.8 but 6.4 for the growth rate. Assuming similar dependencies, probably a lot of crystals were formed during pH-adjustment and a certain amount already reached the detectable size. A group of crystals, however, which were formed during pH-adjustment could not exceed the detectable size. As the high local supersaturations were quickly degraded upon mixing, their growth was only driven by moderate supersaturations which considerably decelerated the growth rate. This would explain why the increase in crystal concentration from 10^7 to 10^9 crystals/kg was visible after a certain lag phase and why this number seems to have stayed constant for a certain time, the length of which depended on the supersaturation of the well mixed system. The latter may subsequently result in an increase in crystal concentration by homogeneous nucleation. The length of the lag phase reflects the high dependence of the nucleation rate on the supersaturation. While this lag phase was almost indefinable for CP01 (initial supersaturation ratio of 6.6), it approximately lasted six hours for CP02 (initial supersaturation ratio of 5.5) but at least 22 hours for CP03 (initial supersaturation ratio of 3.8).

The considerations only include homogeneous nucleation which is probably oversimplified. Secondary nucleation due to crystal breakage and attrition may be of minor relevance due to the low crystal concentration in early stages of the process. However, heterogeneous nucleation on foreign particles is practically unavoidable and often the dominating nucleation

mode in batch crystallisations (Tavare, 1991). For the experimental conditions employed, it is nevertheless argued that heterogeneous nucleation is of minor relevance. The mother liquor was filtered via 0.22 μm membranes directly before the supersaturation was induced by pH-adjustments. The amount of dust particles inside the Eppendorf tubes in which the crystallisation process was conducted was probably not very high since they were stored in a closed plastic bag only opened shortly prior to experimentation. Moreover, the Eppendorf tubes used were PCR (Polymerase Chain Reaction)-clean which means that they are clean-room manufactured. During the process, the tubes were closed preventing the intrusion of further dust particles. Nevertheless, a certain competition between homogeneous and heterogeneous nucleation must be taken into account. In CP03 the crystal concentration reached a value of 5.35×10^8 crystals/kg which did not considerably change during the following 22 hours (Figure 3.7). Since heterogeneous nucleation becomes more dominant at lower supersaturations, it should have led to a constant increase in crystal number which was obviously not the case which emphasises the assumption of negligible heterogeneous nucleation.

The size distribution of the crystals was narrow and mono-modal which may be interpreted as a sign that the crystals were grown on very similar mechanisms, thus emphasising the dominant role of homogeneous nucleation.

3.6 Conclusions

The analysis of batch crystallisation processes of a recombinant *Bacillus halmapalus* α-amylase conducted in 1.5 mL Eppendorf tubes showed very fast crystal growth but only to a relatively small size of approximately 5 μm in mean diameter, independent of the initial supersaturation ratio the crystallisation was started from, i.e. between 5.5 and 6.6 mg/mL. The size distribution did not significantly change during the process. Secondary nucleation due to crystal breakage or attrition cannot be ruled but probably of minor degree since the crystals encountered at the end of the process were of high integrity. Due to the rapid growth and subsequent cessation together with only weak indications of attrition and crystal breakage it is suggested that the crystals formed can be treated as inert particles which do not act as nuclei. Based on that assumption and on the development of the crystal concentration it is proposed to subdivide the crystallisation process into two different phases: The first phase consists of the instant formation of approximately 10^7 crystals/kg caused by high local supersaturations generated for a short period upon induction of the supersaturation by pH-adjustment. Homogeneous nucleation is probably the dominating mechanism. The increase in crystal concentration from 10^7 to 10^9 crystals/kg is a result of the measuring technique which can only detect particles larger than 2 μm. This increase is still a consequence of the nucleation caused by the pH-adjustment. The crystal growth is significantly decelerated due to the lower supersaturations in the mixed system, thus it takes a longer time until the crystals formed by pH-adjustment have grown to the detectable size. Only now the second phase which is also dominated by homogeneous nucleation starts which lasts until the supersaturation is decomposed. The time needed to reach equilibrium on the supersaturation is very dependent on the initial supersaturation ratio. For 6.8 the process was essentially finished after 22 hours whereas an initial supersaturation of 3.8 was not enough to initiate the second phase of the crystallisation within 27 hours. The strong dependence of the time to reach equilibrium on the initial supersaturation ratio suggests that both the nucleation and the growth rate are also highly dependent on the supersaturation. The approach presented in the current study provides valuable information on the kinetics of the crystallisation process and the crystal size and size distribution to be expect. Since the experiments can be conducted in mL-scale, influences of additives or process parameters like stirring rates or temperature on the crystallisation process can be studied in a protein-saving way. Although nucleation and growth mechanisms may significantly differ in larger scale and thus the validity of the findings generated in mL-scale

have to be critically verified, it is believed that the knowledge obtained by the approach chosen in this study can beneficially be included in the design of reliable bulk protein crystallisation processes leading to crystals of desired properties.

4 Study of *Bacillus licheniformis* α-amylase solubility on either side of its isoelectric point

Cornelius Faber,[a,b] Timothy J. Hobley,[b,*] Jørgen Mollerup,[c] Owen R. T. Thomas,[b,d] and Svend G. Kaasgaard,[a]

[a] Novozymes A/S, Novo Alle, DK-2880 Bagsværd, Denmark

[b] Center for Microbial Biotechnology, BioCentrum-DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[c] Department of Chemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[d] Department of Chemical Engineering, School of Engineering, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

*Corresponding author

e-mail: th@biocentrum.dtu.dk (T.J. Hobley)

Telephone: +45 45 25 27 06

Fax: +45 45 88 41 48

Synopsis: The influence of selected anions and cations from the Hofmeister series on the solubility of *Bacillus licheniformis* α-amylase on either side of its isoelectric point is studied and discussed.

Keywords: anions; cations; crystallisation; Hofmeister series; polydispersity; proteins; salts; zeta potential;

4.1 Abstract

The solubility of recombinant *Bacillus licheniformis* α-amylase (BLA) has been studied at pH values on either side of its isoelectric point (pI) at which the enzyme carries either a net positive or net negative charge. A semi-pure preparation was chosen, containing five active isoforms with pIs ranging from 6 to 7.3, together with small amounts (<1%) of other protein impurities. Solubility curves were determined by batch crystallisation with a series of anions and cations at final concentrations of 0, 0.1 and 0.5 M. The lowest solubility of the salt-free solution was 60 mg/mL at pH 7. The addition of 0.1 M sodium salts of nitrate, sulphate and thiocyanate exerted small effects on solubility compared to the presence of 0.5 M salt, where the solubility was significantly lowered by sodium sulphate at all pH-values, and increased by sodium thiocyanate at pH-values 7 and 8. The effect of anions on α-amylase solubility was observed to follow the Hofmeister series, and no reversal of the order was seen on crossing the isoelectric point. The cations had little effect on solubility. The sign and the magnitude of the α-amylase's zeta potential were determined in the presence and absence of 0.1 M salt. Under all tested conditions, the zeta potential was positive at pH 6 (between 2.5 and 5.5 mV) and negative at pH-values of 7 (between -1 and -3.5 mV) and 8 (between -2.5 and -4 mV). The pH of zero zeta potential (i.e. pI) of BLA was found to be at its lowest (pI 6.5) in the presence of 0.1 M sodium thiocyanate (cf. 0.1 M sodium sulphate, pI 6.7; 0.1 M lithium nitrate, pI 6.7; and salt-free solution, pI 6.7). Qualitatively, measurements of the zeta potential correctly predicted the influences the different tested salts imposed on the BLA's solubility.

4.2 Introduction

Due to the demands for new types of highly concentrated protein formulations of industrial enzymes, as well as alternative methods of delivery within the pharmaceutical industry, interest in solubility and bulk crystallisation of proteins has recently increased (Klyushnichenko, 2003). The rapid development of scalable, reproducible and robust crystallisation processes in product recovery and formulation leading to competitive products requires detailed knowledge of the solid-liquid equilibrium (Agena et al., 1999) together with information on how various precipitants such as salts affect the solubility properties of a given protein, particularly when operating at or above its solubility limit. Comprehensive solubility

data is available for only very few proteins, probably the best example being highly pure lysozyme. Further, the number of studies published on bulk crystallisation in the presence of impurities is very low. Indeed, the crystallisation of ovalbumin (Judge et al., 1995) and nucleation and growth of microbial lipase crystals (Jacobsen et al., 1998) represent two of the very few examples. In contrast, the influence of salts on the solubility of highly purified proteins, such as lysozyme, collagenase and carboxyhemoglobin, has been intensively studied (Cacioppo & Pusey, 1991b; Carbonnaux et al., 1995; Green, 1932; Guilloteau et al., 1992; Hofmeister, 1888). The strong impact of salts on a protein's solubility has been demonstrated in all cases and a characteristic order of efficiency has been found, in which the salts, i.e. cations and anions, precipitate the protein out of solution. This order of efficiency is commonly referred to as the Hofmeister series (Hofmeister, 1888) and has been confirmed for many, but not all proteins, e.g., reversal of the anion series has been observed for lysozyme (Riès-Kautt & Ducruix, 1989). Riès-Kautt & Ducruix (1989) argue that the order of this series is dependent on the sign of the protein's net charge and should thus change should the pH of the solution be shifted from values less than to values more than the isoelectric point (pI) of the protein, or vice versa. A reversal of the order of efficiency in which anions influence the solubility depending on the protein's net charge has only been shown for fibrinogen (Leavis & Rothstein, 1974), and while this reversal should not occur for cations (Riès-Kautt & Ducruix, 1997) this has never been systematically verified. From a process point of view, an inverse of the effects of different salts on solubility depending on the sign of a protein's net charge would be of significant importance in the case where salts are used to steer the protein's solubility. Accordingly, the aim of this study has been to further the understanding of how salts influence the solubility of proteins, and evaluate whether the conclusions made by others for highly purified protein systems (e.g., lysozyme) can be transferred to the less pure protein characterised here. Of particular interest to us was how anions and cations influence the solubility of a protein depending on the sign of the net charge. Evaluation of zeta potential measurements for estimating the sign and magnitude of the net charge in impure protein solutions coupled with its utility in aiding rapid identification of conditions of low and high solubility was a secondary objective of this study. A recombinant *Bacillus licheniformis* α-amylase (BLA) was selected as a model protein because it crystallises readily and offered the possibility to conduct solubility studies at pH-values on either side of its isoelectric point.

4.3 Materials and methods

4.3.1 Preparation of *Bacillus licheniformis* α-amylase

Recombinant *Bacillus licheniformis* α-amylase was expressed in *Bacillus licheniformis* fermentation and purified at Novozymes A/S, Bagsværd, Denmark. Cells and other debris were removed by filtration, and the filtrate was concentrated by ultrafiltration (10 kDa cut off). The enzyme concentrate was crystallised by lowering the pH to 7.5, and the crystals subsequently harvested by centrifugation formed the initial feedstock employed in this study.

4.3.2 Reagents

Sodium nitrate was obtained from Merck (Darmstadt, Germany). Sodium sulphate and calcium chloride were purchased from J.T. Baker (Deventer, The Netherlands) and sodium thiocyanate, lithium nitrate and caesium nitrate were supplied by Sigma-Aldrich (Steinheim, Germany). The buffers, 2-(N-Morpholino)ethanesulphonic acid (MES) and 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) were also obtained from Sigma-Aldrich, whereas boric acid was supplied by AppliChem, (Darmstadt, Germany). Sodium hydroxide and acetic acid (at concentrations of 0.1 and 1 M, respectively) were provided by Bie & Berntsen (Rødovre, Denmark). All chemicals were of analytical grade.

4.3.3 Crystallisation studies

All crystallisation processes were conducted using a buffer composed of 10 mM MES, 10 mM HEPES and 10 mM boric acid. All samples were prepared in ultra-pure water of 18.2 MΩcm resistivity (Millipore, Billerica, MA, USA), together with 0.2 % (w/v) of the antimicrobial agent, Proxel LV (Avecia, Manchester, UK) to prevent any interference from microbial contamination.

Depending on the pH-value of interest, crystals of α-amylase were either (i) dissolved in the buffer at pH 9.5, and then filtered through a 0.22 μm pore size cellulose acetate filter

(Sartorius, Göttingen, Germany) or (ii) suspended at pH 5.5 and stirred until they dissolved; subsequently more crystals were added until no further material could be dissolved. The resulting crystal slurry was centrifuged and the supernatant filtered through a depth filter (Seitz, Bad Kreuznach, Germany) of 0.22 μm pore size. The crystallisation experiments were conducted in batch mode and for every condition tested, solutions of two different initial protein concentrations were prepared in the buffer as specified above, containing the precipitants of interest (i.e. salts) as required. With the exception of lithium nitrate, which was added in liquid form as a 3 M stock solution prepared in the buffer described above, all other salts were added in solid form. For each preparation, the crystallisation was initialised by slowly (over 5 minutes) raising or reducing the pH to the value of interest, through the addition of sodium hydroxide or acetic acid respectively. Aliquots (750 μL) of the solution were then transferred into a series of 1.5 mL Eppendorf tubes before placing in a thermo mixer (Model 5355, Eppendorf, Hamburg, Germany), which ensured precise temperature control (40 ± 1°C) and appropriate mixing (1400 rpm). On the basis of visual inspection, it was apparent that no crystals or precipitates formed during pH-adjustment and liquid handling. After 96 hours of incubation, the tubes were removed from the mixer and the precipitates formed were analysed by light microscopy to determine whether they were amorphous or crystalline. In cases where the precipitate was found to be amorphous, the experiment was repeated using a lower initial protein concentration. Crystals were separated from the liquid by centrifugation at 25000 g_{av} for 2 minutes in a temperature-controlled microcentrifuge (Model 5415 R, Eppendorf) operated at 40°C. The supernatants were collected and filtered through 0.22 μm filters (Sartorius) and kept at -30°C until required for analysis. Upon thawing, no crystals were observed in any samples. The pH was checked at the end of a given crystallisation process, and no adjustments were made during an experiment. Deviations from the desired pH value did not exceed ± 0.2 pH-units. In a separate experiment, the supernatant protein concentration was determined at different experimental times. For all tested conditions, the protein concentrations in the supernatant did not change significantly (i.e. < 3%) between 72 and 96 hours, indicating that the equilibrium concentration had to a large extent been approached within the experimental time employed in the current study.

4.3.4 Analysis

The concentration of total protein was measured using a Cobas Fara spectrophotometric robot (Roche, Rotkreutz, Switzerland) with the ESL assay (Roche, Mannheim, Germany), which is based on a reverse biuret method combined with a copper-bathcuproine chelate reaction (Matsushita et al., 1993). The Cobas Fara was programmed to preheat the reagents and samples for 10 minutes at 25°C, then mix them and 30 seconds later measure the absorbance at 485 nm. Protein concentrations were determined from a standard curve constructed using bovine serum albumin (BSA), and are expressed in BSA equivalents The salts employed in crystallisation studies were shown not to interfere in the assay.

The crystals in the feedstock were re-dissolved and checked for purity by reducing SDS-PAGE (NuPage 10% Bis-Tris-gel NP0301, Invitrogen, Carlsbad, CA, USA). A standard mixture of proteins of known molecular weight (LMW 17-0446-01, GE Healthcare, Uppsala, Sweden) was used to characterise the proteins encountered in the bands. Isoelectric points of BLA were determined by isoelectric focusing (IEF) in 10-well Novex gels with a pH-range of 3 to 10 (Invitrogen), and the gel was calibrated using a commercially obtained protein mixture (SERVA liquid mix IEF markers pH 3-10, Invitrogen). IEF runs were conducted in an XCell Surelock Mini-Cell using Novex IEF anode and cathode buffers pHs 3 to 10 (Invitrogen) according to the manufacturer's instructions. All gels from SDS-PAGE and IEF runs were stained with Coomassie Brilliant Blue R-250 (Sigma-Aldrich, Steinheim, Germany). Relative estimates of contaminating proteins and isoform compositions in SDS-PAGE and IEF gels respectively were obtained by Scanning Densitometry using the gel analysis software Quantity One (Bio-Rad, Hercules, CA, USA).

Amylase activity in the IEF gels was confirmed in the following way. Bands were cut from an unstained gel, using a stained gel as template. The gel fragments were placed in 600 μL Eppendorf tubes that had been pierced at their base with a fine gauge syringe needle. These tubes were then placed in 1.5 mL Eppendorf tubes and spun in a microfuge to homogenise the gel fragments. The homogenised gel slurries now contained in the 1.5 mL Eppendorf tubes, were resuspended with 1 mL of 15 mM $CaCl_2$. After incubation for 1 hour, the samples were re-centrifuged to remove the extracted gel pieces. The presence of amylase activity in the supernatants was qualitatively demonstrated using the Phadebas kit (Product number 10-5380-33, GE Healthcare).

The supernatant obtained from batch crystallisation experiments was diluted 10-fold using buffer of the same composition (i.e. with respect to salt type, concentration and pH) as that of the samples. The samples were filtered through 0.22 μm filters (Sartorius) and the zeta potential of each was measured in a Zetasizer Nano ZS (Malvern Instruments, Malvern, UK) using folded capillary cells of 0.75 mL sample volume (Malvern Instruments). Phase analysis light scattering (PALS) was employed to determine the electrophoretic mobility from which the zeta potential was calculated using the Smoluchowski approximation. As the protein and salt concentrations) were low, i.e. < 10 mg/mL and 0.1 M respectively, the viscosity of the samples was assumed to be equal to that of water. Measurements were conducted in triplicates, each of which consisted of 30 individual measurements, and the standard deviation of the mean zeta potential values was in the range of ± 1 mV. The zeta potential was shown to be independent of the dilutions employed in the current study, as has also been reported for other systems, e.g., subtilisin (Pearson et al., 2004). Zeta potential measurements were only conducted at 0.1 M salt concentrations; at 0.5 M the data quality deteriorated significantly, and reproducible values of the zeta potential could not be obtained. No general rules as to the maximum salt concentration which permit useful determinations of the zeta-potential to be made can be given, and should be determined carefully for each system and device being tested. Successful measurements of the zeta potential of proteins (e.g., bovine serum albumin at 0.1 M sodium chloride (Mattison & Kaszuba, 2004), *Bacillus halmapalus* α-amylase at a conductivity of 25 mS/cm, Chapter 2), intact and disrupted *E. coli* and yeast cells at 32 mS/cm conductivity (Lin et al., 2003)) conducted with similar equipment have been reported previously. Measurement of the zeta potential of lysozyme crystals at a conductivity as high as 50 mS/cm using monovalent electrolytes has been performed with different instruments. In these studies problems were encountered at certain pH-values, and additionally measurements in divalent electrolytes (i.e. either divalent cations or anions) could not made due to low electric fields (Lee et al., 2001). Reliable zeta potential measurements were achieved in the present study by ensuring that the sample conductivity was always less than the sample conductivity ties 16 mS/cm.

4.4 Results and discussion

4.4.1 Properties of the feedstock

The purity of the feedstock containing BLA was examined by SDS-PAGE. One major band with a size of 55 kDa was observed in Coomassie Blue stained gels, consistent with the molecular weight of the amylase (Figure 4.1 A). A number of species with lower molecular weights were also detected, but these accounted for less than 1% of the total protein as determined by gel densitometry. Isoelectric focusing followed by Coomassie Blue staining revealed five bands on gels with pIs ranging from 6.0 to 7.3 (Figure 4.1 B). The most abundant isoform (pI 6.8) accounted for 30% of the total protein (determined by gel densitometry), and the least abundant (pI 6) for only 10% of the total protein. Interestingly, three isoforms were found in very close proximity to each other in IEF gels, running at pHs 6, 6.15 and 6.3. Collectively, these three species accounted for 45% of the total protein. Amylase activity was associated with all 5 Coomassie Blue stained isoforms (i.e. detected in all five gel slices fragments), although it is not known how they differ from one another. Given that all isoforms ran as a 55 kDa species in SDS-PAGE and that they all exhibit α-amylase activity, it is probable that differences in their pI may be due to modification of certain amino acid side chains, e.g. deamidation. BLA-solutions exhibited a brownish colour, possibly caused by the presence of non-proteinaceous, e.g., small amounts of carbohydrates and salts Although the disturbing influence of contaminants on crystallisation processes has been demonstrated previously (Giegé et al., 1986), the enzyme employed in this study was found to crystallise readily on either side of its pI, forming thin, rhombus shaped crystals. When measured in the buffer specified above, in the absence of salts the solubility of BLA-solutions was lowest at pH 7 with 60 mg/mL protein remaining in solution. At pH 6, the solubility doubled to 125 mg/mL, and at pH 8, approximately 85 mg/mL of BLA could be solubilised (Figure 4.2).

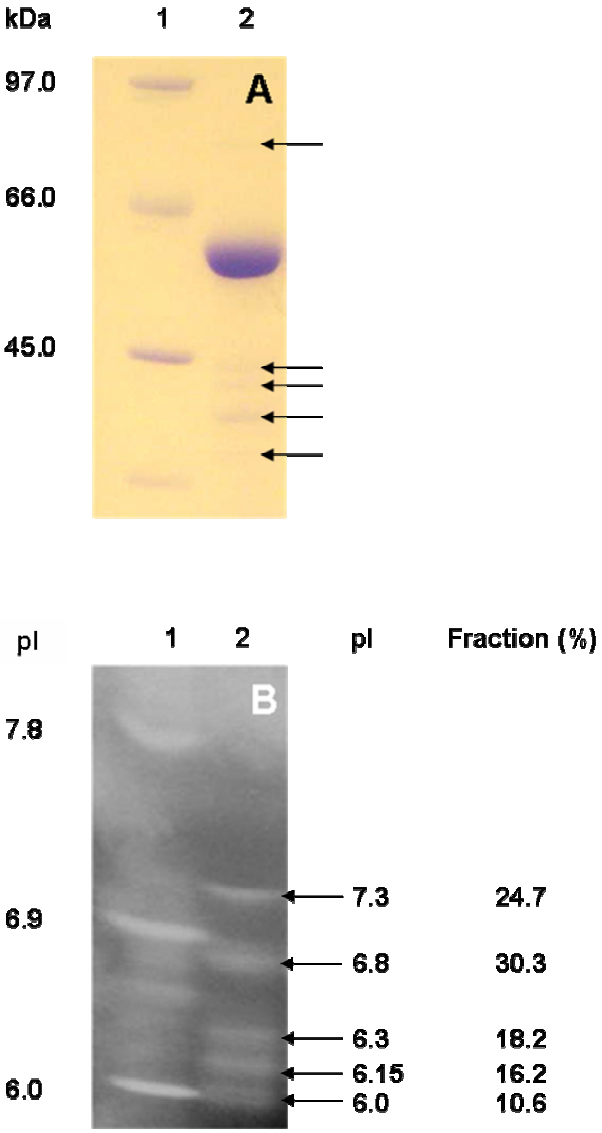

Figure 4.1 **Examination of *B. licheniformis* α-amylase feedstock. (A) SDS-PAGE analysis. The molecular weight standards are shown in lane 1 and the arrows show the position of the bands stemming from contaminating proteins. (B) Isoelectric focusing. The calibration standards are in lane 1. The arrows show the position of the isoforms.**

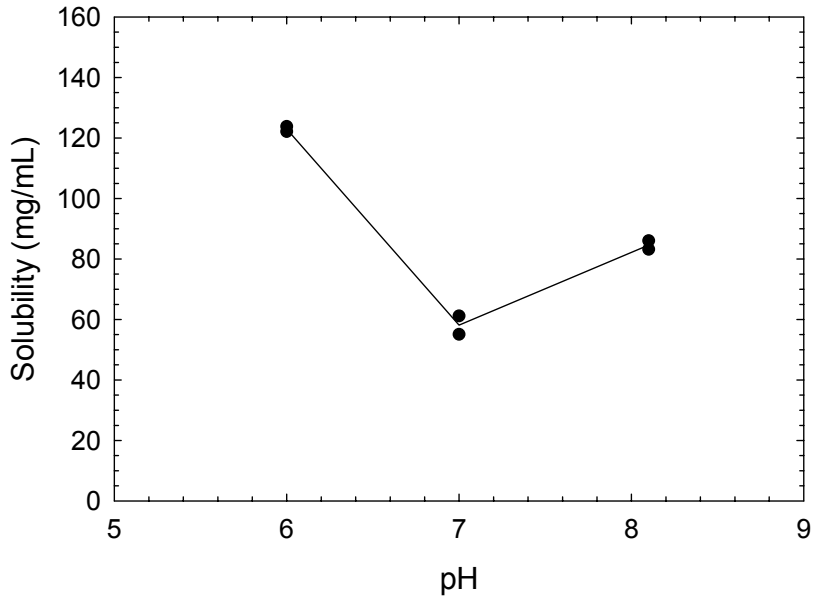

Figure 4.2 BLA solubility as a function of pH. The crystallisation process was conducted in batch mode started from two different initial supersaturations in a MES, HEPES and boric acid buffer (containing 10 mM of each). The experiments were conducted at 40°C for 96 hours.

4.4.2 Solubility as a function of salt concentration

Effect of anions on solubility of BLA

The effects of the sodium salts of nitrate, sulphate and thiocyanate employed at two different concentrations (i.e. 0.1 and 0.5 M) and at three different pH-values (6.0, 7.0 to 8.1) spanning BLA's pI, were examined (Figure 4.3 A and B, respectively). At both concentrations studied, consistent with the Hofmeister series, the sulphate anion resulted in the lowest solubility, followed by nitrate and then thiocyanate. Moreover, this series appeared to be preserved at all pH-values studied, i.e. at the pI and on either side of it (Figure 4.3 A). That said, at pH 6 and 0.1 M salt, the measured solubilities were very similar to one another, and a characteristic order was thus difficult to determine. In the presence of all salts at both concentrations, the solubility at pH 6 was 10 to 20% lower than without added salts (see Figure 4.2), with 0.5 M sulphate reducing the solubility by approximately 35% (Figure 4.3 B). Interestingly, at pH 7 the solubility was increased in the presence of 0.5 M nitrate and particularly 0.5 M thiocyanate, cf. the salt-free case (Figure 4.2), whereas 0.5 M sulphate lowered the solubility

drastically (Figure 4.3 B). These trends were also observed at pH 8 in the presence of 0.5 M of each salt, and to a lesser extent with 0.1 M salt.

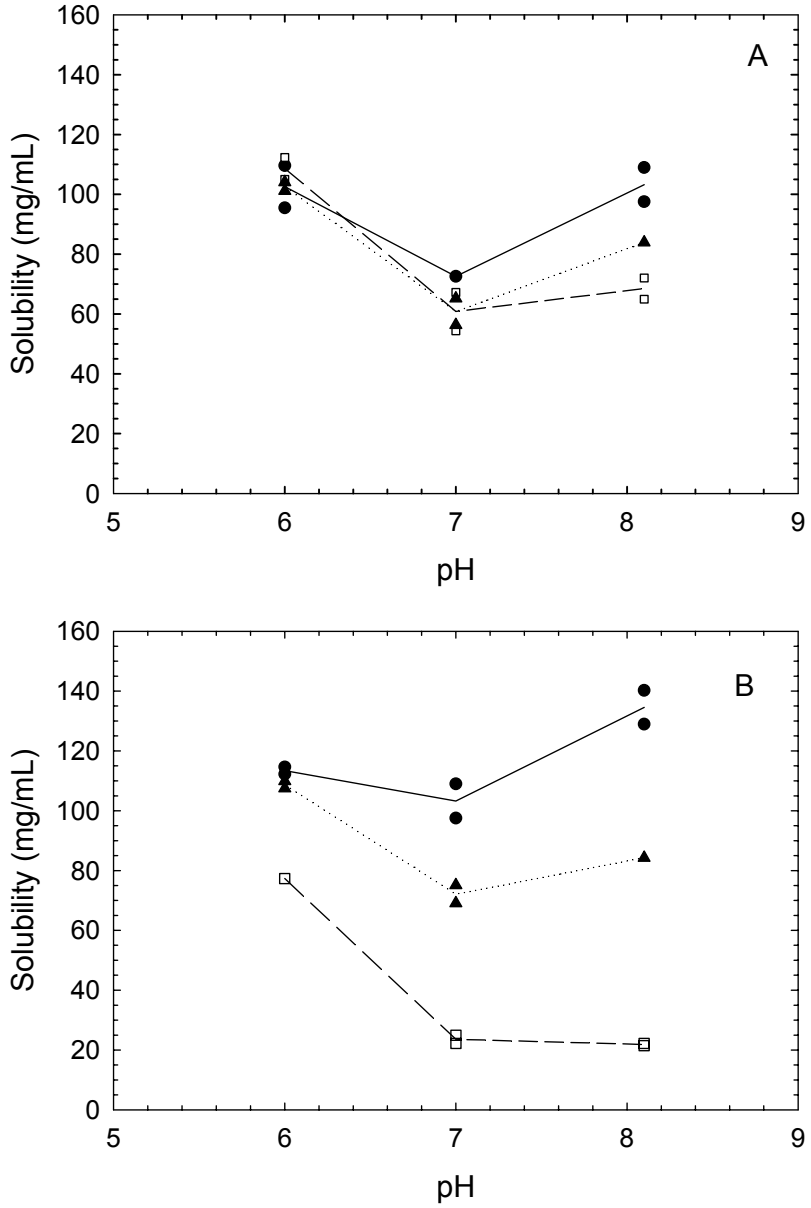

Figure 4.3 Influence of anions on BLA solubility as a function of pH. Influence of sodium thiocyanate (●, solid line), sodium sulphate (□, dashed line) and sodium nitrate (▲, dotted line) at 40°C after 96 h. Crystallisations were started from two different initial supersaturations in a MES, HEPES, boric acid buffer (10 mM of each). Salt concentrations of 0.1 M (A) and 0.5 M (B) were employed.

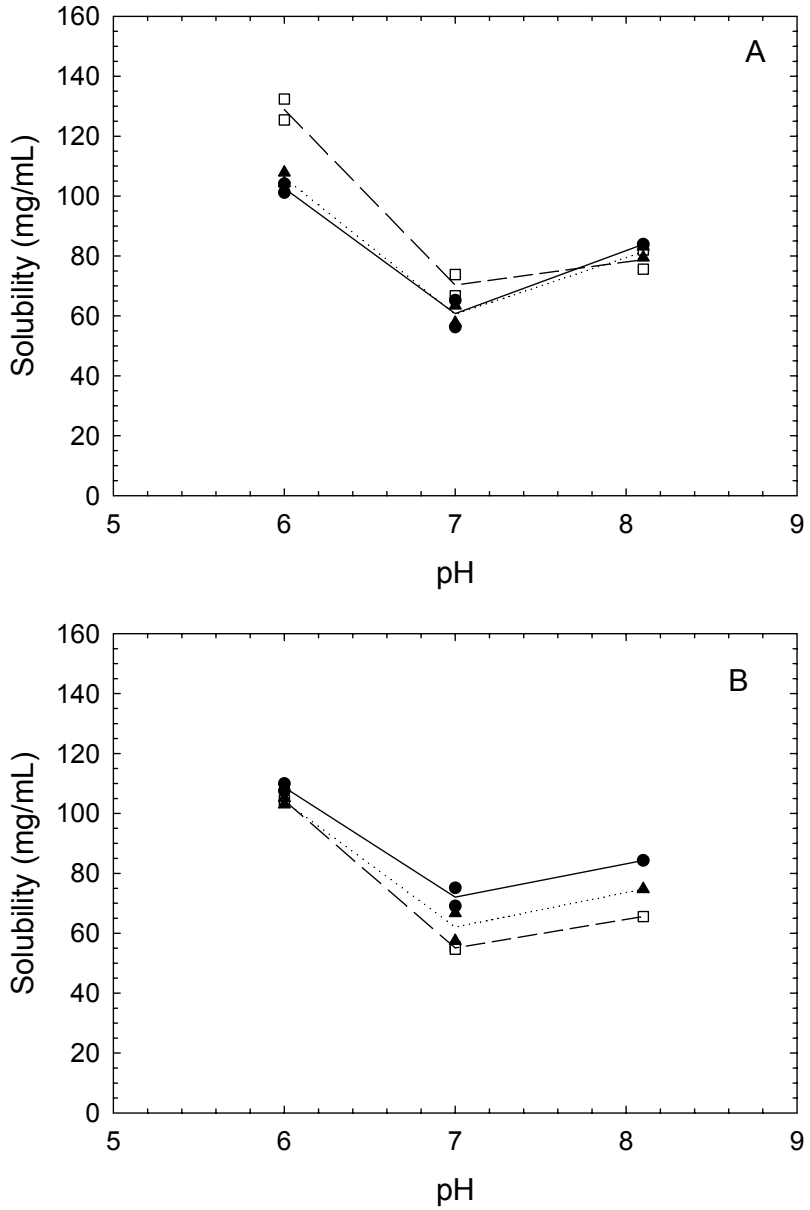

Figure 4.4 **Influence of cations on the BLA solubility as a function of pH. Influence of sodium nitrate (●, solid line), lithium nitrate (□, dashed line) and caesium nitrate (▲, dotted line) measured at 40°C after 96 h. Two different initial supersaturations buffered in MES, HEPES and boric acid (10 mM of each) were used. Salt concentrations of 0.1 M (A) and 0.5 M (B) were employed.**

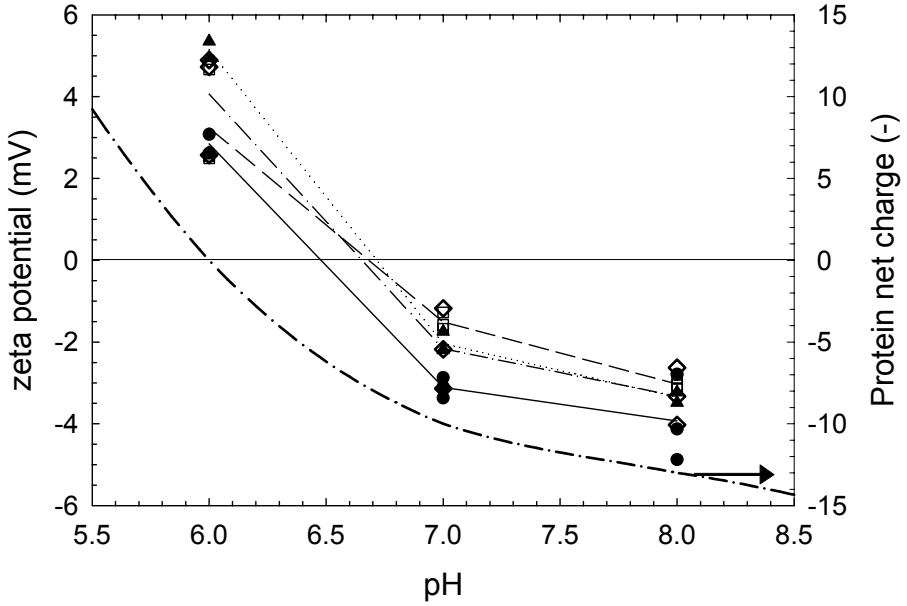

Figure 4.5. Influence of salts on the zeta potential of BLA as a function of pH. Influence of sodium thiocyanate (●, solid line), sodium sulphate (□, dashed line) and lithium nitrate (▲, dotted line) at 0.1 M salt concentration buffered in MES, HEPES and boric acid, 10 mM each. For comparison, zeta potential of the salt-free BLA-solution (◇, dash-dotted line), also buffered in MES, HEPES and boric acid, 10 mM each and the calculated charge curve based on the amino acid sequence of *Bacillus licheniformis* α-amylase is included in the graph (dash-dotted line).

Effects of cations on solubility

The results presented in Figure 4.4 A and B show that the effects of the cations on solubility were much less dramatic than was observed for the anions. In all cases, the solubility was lowest at pH 7 and highest at pH 6. When 0.1 M lithium salt was used, essentially no effect on the solubility at pH 6 was seen compared to the case without any added salt, whereas for caesium and sodium nitrates a 20% reduction at pH 6 occurred. At pH 7, 0.1 M lithium nitrate increased the solubility whereas 0.1 M levels of the other two salts did not impose any notable effects cf. the salt-free control case (Figure 4.2). At pH 8, the addition of the three different nitrate salts at a concentration of 0.1 M exerted hardly any effect on BLA's solubility (Figure 4.4 A), but when added at the 0.5 M level (Figure 4.4 B), some effects were noted. For example, the lowest solubility was observed with lithium followed by caesium and then

sodium. This sequence does not concur with other reports dealing with the influence of cations on solubility. For example, in the case of lysozyme it was reported that lithium was the best precipitant, followed by sodium and then caesium (Riès-Kautt & Ducruix, 1997). In this study at 0.5 M, all salt types lowered BLA's solubility by approximately 20% at pH 6, and by 10 to 20% for caesium and lithium at pH 8 cf. the salt-free case (Figure 4.2). At pH 7 however, the addition of lithium nitrate led to a very slight reduction in solubility, whereas sodium conversely increased the solubility (Figure 4.4 B) to a small extent cf. the case without added salts. This said, the differences between the three nitrate salts at pH 7 were minimal and within the data spread, and we conclude that, the influence of the tested cations on BLA's solubility is small.

4.4.3 Effects of salts on the zeta potential of BLA

In Figure 4.5 the zeta potentials measured for BLA-solutions without added salts and with 0.1 M sodium sulphate, sodium thiocyanate and lithium nitrate are shown alongside the theoretical charge curve for BLA calculated purely on the basis of its primary amino acid sequence. In all of the cases examined zeta potential measurements confirmed that BLA carried a net positive charge at pH 6 and net negative charge at pHs 7 and 8 (Figure 4.5). At pH 6 the addition of 0.1 M sodium thiocyanate resulted in the lowest charge on BLA followed by sodium sulphate, and then lithium nitrate. At pHs 7 and 8, the lowest net charge on BLA was observed in the salt-free case. The presence of thiocyanate yielded the most highly charged protein, followed by lithium nitrate and then sodium sulphate (Figure 4.5). At pH 6 lithium nitrate gave the most charged protein. Although we have not examined the effects of sodium nitrate we anticipate the behaviour of this salt, given our earlier demonstration that the cation plays a relatively minor role in determining BLA's solubility. Thus, we expect a general shift in the counter ions associating with the protein will occur, from the cation sodium at pH 8 and 7 to the different anions employed, i.e. thiocyanate, nitrate and sulphate at pH 6. It is interesting to note that these trends mirror the above solubility data (for 0.1 M salts) exactly (Figure 4.3 A). In both BLA's solubility at pHs 7 and 8 in the presence of thiocyanate is much higher than in the presence of nitrate, which in turn is higher than sulphate. At pH 6 there was no significant difference in BLA's solubility for the three tested salts, as is reflected by the zeta potential measurements in Figure 4.5. Had it been possible to

measure zeta potentials on BLA solutions supplemented with 0.5 M salts, we would expect a similar qualitative trend correlating BLA's charge with its solubility.

From the charge curve data in Figure 4.5 the expected pI for BLA was approximately 6, whereas the pI derived from zeta potential measurements in the absence of added salt was 6.7. The reader should recall that isoelectric focusing (Figure 4.1 B) revealed the presence of five isoforms with pIs ranging from 6 to 7.3. Interestingly, by considering the weight contributions of each band as determined by gel densitometry, we estimate an average pI of 6.6 which agrees surprisingly well with the pI of 6.7 determined from zeta potential measurements conducted on the salt-free BLA solution. It is not uncommon for charge curves based on the amino acid sequence to deviate from experimentally obtained ones (Windsor, 2003). The main source of such discrepancies is that the three-dimensional structure of the protein in question cannot be taken into account when the charge curve is calculated from the amino acid sequence. Moreover, as sequential, structural and chemical differences within the encountered isoforms are not known, their individual influence cannot be considered in the calculations. The zeta potential is based on the average electrophoretic mobility of all protein molecules in solution and leads to one average pH of zero zeta potential, even in the presence of a mixture of proteins of very different structure and pI. Such an average pI is not straightforward to determine by other methods like isoelectric focusing, particularly when the influences of various salts are to be studied. However, such average pI-values are perhaps more relevant for solubility measurements given that solubility is a result of all forces interacting within all the substances present in a particular solution (Riès-Kautt & Ducruix, 1997). Zeta potential and charge should, however respond in a similar way to changes in pH, as Figure 4.5 confirms. Within the pH-range of 0.5 units above and below the pI, both the calculated protein net charge and the zeta potential change relatively rapidly, i.e. both units change from positive to negative values with increasing pH, whereas within the range of between 0.5 and 1.5 pH-units above the pI, less pronounced changes are suggested for both units. It is interesting to note that if the whole charge curve were shifted by 0.5 - 0.7 pH-units in the alkaline directions and appropriately scaled, it would almost exactly superimpose on the experimentally determined zeta potential curve for the salt-free BLA solution (Figure 4.5). In summary, the qualitative agreement between the two curves is taken as evidence for the high accuracy of the zeta-potential measurements.

As mentioned above, measurements of the zeta potential suggested that sodium thiocyanate was capable of changing the pI of BLA towards lower pH-values (i.e. from 6.7 to 6.5). Shifts to lower pH-values indicate anion binding whereas the opposite is true for cations (Mullet et al., 1997). Thiocyanate binding to BLA can therefore be assumed, and is in keeping with reports that the thiocyanate ion binds more tightly and in greater numbers to protein molecules compared to chloride ions, regardless of the sign of the protein net charge (Scatchard et al., 1950).

The reason why the zeta potential of BLA was measured using lithium nitrate instead of sodium nitrate was that we found that 0.2 M lithium nitrate increased the pI of a structurally related amylase (with approximately 85% sequence homology and of comparable purity) by approximately 0.7 pH units (Chapter 2). We attributed this finding to strong binding of the lithium ion to oppositely charged amino acid residues at the protein surface. In stark contrast, in the present study, substantial changes in BLA's solubility (Figure 4.3 and Figure 4.4) and pI (Figure 4.5) were not observed in the presence of lithium. The effect the different ions induce on the solution properties thus seems to be protein specific, and must inspected carefully for each system being studied.

The zeta potential is often used to determine the pI of macromolecules or particulate systems, for characterisation of the binding behaviour of additives, and to estimate the colloidal stability of a system of interest. Solutions which are characterised by zeta potentials >30 mV (either positive or negative) are generally regarded as stable. Below this value phase separation processes are likely to occur (He et al., 2003). The zeta potentials measured for BLA were <6 mV, which is much lower than the proposed threshold of 30 mV, indicating an enhanced likelihood of precipitation processes. Indeed, the BLA system examined is very prone to precipitate formation, either in amorphous or preferentially in crystalline form. Although we have not examined this experimentally, we believe it will be difficult to reach zeta potentials >30 mV, given the limited colloidal stability of BLA between pHs 6 and 11. Carbohydrates and/or higher alcohols may prove to be promising additives in enabling this threshold to be reached thereby improving colloidal stability. However, because the zeta potential is independent of protein concentration (Pearson et al., 2004), it cannot indicate conditions in which the solution is supersaturated, which is a necessary requirement for predicting the outcome of crystallisation processes. In this context, measurements of zeta

potential can only therefore be employed to identify process conditions with increased or reduced risk of phase separation.

In the introduction, it was discussed that previous workers have proposed that the order of the influence of anions on protein solubility is reversed when crossing the pI to change the net charge on a protein, and that this order reversal is a generic phenomenon (Leavis & Rothstein, 1974; Riès-Kautt & Ducruix, 1989). However, the current study did not deliver clear evidence of such a phenomenon with the BLA preparation and salts used herein. Thus in the following discussion the reasons for the behaviour observed are examined. The results show that electrostatic interactions are involved in the solubility behaviour seen, since the pH giving the lowest solubility was in good agreement with the pH of zero zeta potential. The data indicates that electrostatic effects dominate at least up to a salt concentration of 0.1 M and probably still at 0.5 M (Retailleau et al., 2002). Electrostatic interactions of charged groups on the protein molecule with its surroundings vary with the pH of the solution. The results also suggest that the presence of salts can influence both the pI as well as the number of charged groups due to ion binding to appropriate residues on the protein surface and due to changes of the dissociation constants of acidic or alkaline protein groups.

A number of models have been proposed to explain solubility by electrostatic interactions and the Debye-Hückel-theory (Debye & Hückel, 1923) and the Kirkwood model (Kirkwood, 1934) were the first theoretical approaches. However, the validity of the models is restricted to proteins carrying zero net charge (Retailleau et al., 1997) and the influence of impurities on the protein solubility can not be considered. Consequently, many exceptions from the predictions of the two models have been reported (Bénas et al., 2002; Riès-Kautt & Ducruix, 1989; Riès-Kautt & Ducruix, 1991). Approaches of broader validity are to describe the capability of salts to modify solubility by: (i) altering the surface charges on a protein, or (ii) to perturb the interactions between the protein and the bulk solvent water (preferential hydration) (Weber, 1991). The interplay between the two different approaches is believed to have great importance for the solubility upon salt addition and may be the key to whether to expect a reversal of the Hofmeister series or not.

In the first approach, salts can modify the protein surface charge using two mechanisms, either by charge screening or by ion pairing. In both cases, salt ions approach oppositely charged amino acid residues of a given protein. In the case of charge screening, the ion

associates with the residue whereas in ion pairing it covalently binds to the residue. The two mechanisms differ mainly in the binding affinity between ion and residue whereas the effect on the protein surface charge is comparable (Retailleau et al., 2002). The degree of charge screening or ion binding to the surface of macromolecules may be estimated by the interaction of the ion with the surrounding water molecules and the ability of the ion to penetrate the ordered water layers which may be associated with the protein binding sites. Highly hydrated ions such as lithium and sulphate (kosmotropes) do not readily form ion pairs as the surrounding ordered water layer has a shielding effect on the electrical potential of the ion. In contrast, less hydrated ions such as thiocyanate and nitrate are not shielded and tend to even disrupt the surrounding water structure (chaotropes), so that their access to the protein molecule is improved (Jelesarov et al., 1998).

In the second approach, salts have an influence on the interaction between the protein and the solvent, here the bulk water. Salts change the chemical potential of the protein in solution by affecting the partition of water between the protein and the salt-ions. Usually, the protein has a higher affinity for water than for the salt which leads to the formation of a precipitant-depleted zone (or hydration layer) near the protein surface (preferential hydration) which is thermodynamically unfavourable. To decrease the area of the precipitant-depleted zone, the protein molecules associate which leads to a decrease in solubility (Timasheff & Arakawa, 1988). The degree of the preferential hydration is steered by the characteristics of the salt ions involved. Kosmotropic ions like sulphate are more likely to be excluded from the protein surface so that they are more effective in inducing the association of protein molecules than chaotropes.

From the above, it is only the first approach, of charge screening and ion pairing that is dependent on the protein net charge and these are thus the only mechanisms which could lead to a reversal of the Hofmeister series. The number of charged amino acids on the protein surface is believed to be important for how sensitive the solubility is to inverting the sign of the proteins net charge. In light of this, the following can be concluded from the solubility studies presented in the current work: At negative protein net charges, i.e. pH 7 and 8, the increase in solubility caused by thiocyanate and to a lesser degree by nitrate was caused by charge screening or ion pairing (Figure 4.3). Charge screening by the sodium ions also has to be considered (Figure 4.4 A and B) but seemed less effective compared to thiocyanate binding since the solubility increased (Figure 4.3) but almost equally effective as nitrate binding or

screening. In contrast, the decrease in solubility observed for sulphate is believed to be a result of the second approach of preferential hydration of the protein molecule, thus causing protein association. At the same time, due to the divalence of sulphate, two-times as many sodium ions were present compared to nitrate and thiocyanate, and thus could have screened negative charges of the protein. The latter would however, require that sodium ions could penetrate the precipitant-depleted layer to gain access to the negative patches. This study cannot conclusively indicate whether charge screening and ion pairing were dominant or whether preferential hydration led to a reduction in solubility. However, since cations play a minor role on protein solubility (Green, 1932; Hofmeister, 1888), preferential hydration caused by sulphate is believed to be the prevailing mechanism.

At positive net charges of the protein, a decrease in BLA solubility was observed with all salts tested, which in the case of thiocyanate and nitrate could be the result of preferential hydration as well by charge screening or ion pairing. As discussed above, sulphate is believed to mainly act by preferential hydration. It is noteworthy that at 0.1 M salt concentration, the tested anions had a very similar effect on the solubility whereas at 0.5 M, sulphate lowered the solubility significantly more. At positive protein net charges, the effect of the sodium ions is believed to be negligible (Riès-Kautt & Ducruix, 1991). Preferential hydration is probably enhanced at increasing concentrations of kosmotropic ions like sulphate, which would thus explain why the salt caused such a strong decrease in solubility compared to the chaotropes thiocyanate and nitrate.

The results presented may be suited for an estimation of whether the two mechanisms by which thiocyanate and nitrate influence the solubility (i.e. charge screening or ion pairing and preferential hydration) act synergistically. The results, however, suggest that the mechanisms almost counterbalance each other. Preferential hydration, i.e. the formation of a precipitant-depleted layer around the protein, limits access to the protein surface, which hinders counterions from screening or binding to amino acid residues. Since nitrate and thiocyanate are chaotropic ions, the effect of preferential hydration is probably less strong compared to sulphate which means that the access of nitrate and thiocyanate ions to the protein surface is still given. X-ray studies conducted on positively charged lysozyme confirmed penetration of the hydration layer by counterions such as thiocyanate or nitrate (Vaney et al., 2001). This could lead to a local disruption of the hydration layer which would lower the propensity of the protein molecules to associate with each other, thus leading to a less strongly reduced

solubility compared to a situation in which either one or the other mechanism is clearly dominating.

The effect of charge screening and ion pairing on solubility is proposed to be stronger at increased protein net charge and at a certain point to become dominant over effects caused by preferential hydration. Thus if charge screening or ion pairing dominated, a reverse of the Hofmeister series could indeed be expected. It is argued that the interplay between the two effects is protein and salt specific and may, among other parameters, depend on the size of the protein molecule, on the number of polar amino acid residues exposed to the surface and on the character of the cations and anions involved (i.e. whether they are chaotropic or kosmotropic). The general assumptions for anions should also be valid for cations.

In light of the above, we speculate that the reasons for the absence of a reversal of the Hofmeister series in this work result from the presence of small amounts of impurities such as carbohydrates and contaminating proteins. Moreover, the isoforms present significantly differ in their pI, ranging from 6 to 7.3. Although measurements of the zeta potential and solubility suggested an overall pI of the solution significantly higher than 6 (i.e. 6.7), the three isoforms of pI-values of between 6 and 6.3 which together account for approximately 45% are believed to play a crucial role in preventing the expected reverse. Had it been possible to measure the solubility and the zeta potential at pH-values significantly below 6, a reverse in the order of both anions and cations would be expected.

4.5 Conclusions

The current study demonstrated that concepts developed on highly purified proteins are only partly applicable to systems of minor purity such as BLA. In particular, no evidence for the reversal of the Hofmeister series for anions and cations upon changing polarity of the protein net charge was found. Deviations from the solubility behaviour predicted for highly purified proteins, e.g., upon salt addition, should thus be expected. Although tedious, solubility measurements are inevitable to correctly characterise the solution properties of the system of interest which is crucial for the development of reliable crystallisation processes for bulk enzyme recovery. The zeta potential has been proven helpful to describe the solution properties of BLA, particularly in the presence of impurities and isoforms. Measurements of

the zeta potential enabled the rapid determination of the protein's pI and could, although limited to low ionic strengths, in qualitative terms correctly predict the impact of the tested salts on solubility. Thus, critical conditions, e.g., of low solubility as well as formulation additives which enhance the colloidal stability of a given protein could rapidly be identified. However, zeta potential measurements cannot replace but help reducing the number of solubility experiments by focusing on critical conditions and thus help accelerating the development of bulk enzyme recovery processes.

5 Strategy for the rapid generation of entire phase diagrams of aqueous solutions of *Bacillus halmapalus* α-amylase using microtitre plates

Cornelius Faber,[a,b] Timothy J. Hobley,[b,*] Jørgen Mollerup,[c] Owen R. T. Thomas,[b,d] and Svend G. Kaasgaard,[a]

[a] Novozymes A/S, Novo Alle, DK-2880 Bagsværd, Denmark

[b] Center for Microbial Biotechnology, BioCentrum-DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[c] Department of Chemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[d] Department of Chemical Engineering, School of Engineering, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

*Corresponding author

e-mail: th@biocentrum.dtu.dk (T. Hobley)

Telephone: +45 45 25 27 06

Fax: +45 45 88 41 48

Synopsis: Phase diagrams of *Bacillus halmapalus* α-amylase under a variety of conditions have been generated using microtitre plates.

Keywords: amorphous precipitation; crystallisation; microtitre plates; nucleation; precipitation; proteins; phase diagrams;

5.1 Abstract

Phase diagrams consisting of precipitation, nucleation, metastable and undersaturated zones under a variety of conditions have been constructed for a recombinant *Bacillus halmapalus* α-amylase (BHA) by means of 96-well microtitre plates incorporating solubility data available from a previous study. Phase diagrams were established as a function of pH in the range of 6 to 10.5, of sodium chloride concentration at pH 7 and 9, and of sodium thiocyanate concentration at pH 7, between 0 and 1 M. Supersaturations were induced by dissolving a large amount of BHA at pH-values giving high solubility followed by adjustment of the pH and salt concentrations to the desired values. The batch phase separation processes conducted at working volumes of 160 μL were finalised after 96 hours and the resulting precipitation inspected for crystalline or amorphous character by light microscopy. After 96 hours, the metastable zone was found to be small within the entire tested pH-range in the presence of the buffer only. The width of the metastable zone could be manipulated by the addition of salt. The biggest increase was observed for sodium thiocyanate and with sodium chloride, the broadening effect was much more pronounced at pH 7 than at pH 9, for the latter the metastable zone was extremely small. In all cases, large nucleation zones were identified and high relative supersaturations ranging from 67 to 110 were necessary to induce the formation of amorphous precipitate. For the tested conditions, no amorphous precipitate was formed above pH 7, regardless of the presence of salt. The borderline between precipitation and the nucleation zone followed the same general trends as the solubility curve, indicating that their position was governed by similar mechanisms. Rod-shaped crystals were found under all tested conditions but their size and their length-width-ratio varied considerably, depending on the supersaturation and salt type employed. Interactions between the different zones of the phase diagram and possible underlying mechanisms are discussed, thereby putting the phase behaviour of BHA in a broader perspective.

5.2 Introduction

The design of efficient downstream processes and their reliable operation are core issues for the economical production of a number of proteins for pharmaceutical or industrial applications. Due to the absence of comprehensive solubility data, in many cases the design of

downstream processes is based on trial-and-error and experience obtained from other proteins which often leads to unsatisfactory results. The development of truly optimised downstream processes would substantially benefit not only from comprehensive solubility data but even more from the availability of phase diagrams (Agena et al., 1999) which consist of (i) precipitation (i.e. the formation of amorphous precipitate), (ii) nucleation (formation of crystals), (iii) metastable (no crystal formation within the experimental time in spite of supersaturation but growth in size of any existing crystals) and (iv) undersaturated zones (protein fully dissolved and will never crystallise) (Chayen, 2005).

Little is known about how the different zones mentioned above interact with each other. A more profound understanding of the mechanisms behind such interactions would pave the way for identifying conditions to conduct crystallisation giving high recovery yields, desired crystal morphology, as well as for avoiding uncontrolled nucleation or undesired amorphous precipitate formation in unit operations (e.g., in filtration or column purification steps). The width of the metastable zone under process relevant conditions is of great interest since solutions in the metastable zone are particularly difficult to handle. Their appearance is the same as of undersaturated solutions but they are supersaturated (Ducruix & Ries-Kautt, 1990), meaning that any kind of particle intrusion or energy input, e.g., from pumps or stirrers, could induce uncontrolled nucleation and lead to product loss (Feigelson, 1988). On the other hand, solutions in the metastable zone are ripe for seeding, i.e. the directed introduction of already formed crystals, preferably but not necessarily of the same material to be processed, to decouple nucleation from crystal growth (Bergfors, 2003). Seeding offers the distinct advantage of enhanced control over crystal size and distribution, which may be beneficial during harvesting and subsequent product formulation and application. Moreover, detailed knowledge of the solubility curve, the width of the metastable zone and of factors influencing the latter could lead to the formulation of enzyme solutions concentrated beyond the solubility limit, yet with an acceptable risk of undesired crystallisation.

Despite the undisputed advantages of phase diagrams, they are essentially unavailable, mainly because their establishment is very time and material consuming and thus precluded for the vast majority of proteins (Basu et al., 2004; Saridakis & Chayen, 2003). Furthermore, they have to be determined for each particular protein of interest since the phase diagram can be changed in an unpredictable way even for very small differences between protein variants, not to mention different protein types. Phase diagrams of *Bacillus halmapalus* α-amylase (BHA)

consisting of precipitation, nucleation, metastable and undersaturated zones as a function of pH and salt type and concentration were generated in this study using microtitre plates which allowed the simultaneous analysis of many conditions for their precipitation properties yet at acceptable protein consumption. Microtitre plates have a wide field of applications in drug testing, genetic studies, and combinatorial chemistry (Kiesslich et al., 2003) and are now a standard in analytical research and clinical diagnostic testing (Kreusch et al., 2003) since they are ideal for high-throughput studies (dos Santos et al., 2002). However, the use of microtitre plates to generate phase diagrams of proteins has to the author's knowledge not been reported before. BHA solubility curves which were available from experiments conducted previously (Chapter 2) were incorporated to complete the phase diagrams. BHA was chosen as a model protein because it is stable between pH 6 and 11 enabling the examination of a large range of different conditions.

5.3 Materials and methods

5.3.1 Preparation of *Bacillus halmapalus* α-amylase

Bacillus licheniformis grown on a complex medium was used to express the recombinant *Bacillus halmapalus* α-amylase, which was produced and purified at Novozymes A/S, Bagsværd, Denmark. Cells, debris and other solids were removed by filtration, and the resulting solution was concentrated by ultrafiltration (10 kDa cut off), and then crystallised by lowering the pH to 7.5. The crystals were harvested by centrifugation and used as feedstock for this work. For all experiments, a buffer containing 10 mM MES, 10 mM HEPES (both from Sigma-Aldrich) and 10 mM boric acid (AppliChem, Darmstadt, Germany), all of analytical grade, was used. These reagents were dissolved in ultra-pure water of 18.2 MΩcm resistivity (Millipore, Bedford, MA, USA). To prevent any disturbance from microbial growth, 0.2 % (w/v) of the antimicrobial agent, Proxel LV (Avecia, Manchester, UK), was added. Crystals of α-amylase were then dissolved in the buffer at pH 11, and after dissolution, the solution was filtered through a 0.22 μm pore size cellulose acetate filter (Sartorius, Göttingen, Germany). The resulting protein concentration was 80 mg/mL.

5.3.2 Reagents

Sodium chloride was purchased from Fluka (Buchs, Switzerland), and sodium thiocyanate from Sigma-Aldrich (Steinheim, Germany). Sodium hydroxide and acetic acid (at concentrations of 0.1 and 1 M respectively) were obtained from Bie & Berntsen (Rødovre, Denmark). All chemicals were of analytical grade.

5.3.3 Sample preparation

The filtered stock solution (see above) was used to generate a dilution series comprising in total eight different protein concentrations. In the case where the phase diagram was to be determined as a function of pH, the pH was lowered stepwise to the desired value in increments of 0.5 pH units by slowly adding 0.1 M or 1 M acetic acid to each sample. The volume of added acid was noted down to be considered in the determination of the protein concentrations. To generate the phase diagram as a function of salt concentration, a BHA-stock solution and a dilution series were prepared as described above together with a 3 M salt stock of pH 11 using the same buffer as specified above. The salt stock was added to each fraction of the dilution series so that salt concentrations between 0 and 1 M were obtained. The pH was lowered to the desired value by adding 0.1 and 0.5 M acetic acid and the added volume noted.

For all experiments, 160 µL of each solution was transferred to 96-well flat bottom microtitre plates (Nalge Nunc International, Rochester, NY, USA). Since the different starting conditions were prepared manually, the difference in time between the first (A2) and the last well (H11), prepared row by row, was between three and four hours. The plates were then sealed with radiation-sterilised adhesive tape (Nalge Nunc International, Rochester, NY, USA) and placed on Eppendorf thermo mixers (Model 5355, Eppendorf, Hamburg, Germany) that permitted precise temperature control (± 1°C). The crystallisation experiments were conducted at 40°C and 600 rpm. Once a day, a picture of the microtitre plates was taken to follow the phase transition process. The experiments were terminated after 96 hours. The pH was not adjusted during the experiments but found to stay constant within a range of ± 0.2 units of the designated value. The precipitation behaviour was analysed by light microscopy (BH 12, Olympus Corp., Tokyo, Japan, equipped with an Olympus DP 12 camera and

interface system) so that areas of precipitation, crystallisation and corresponding crystal morphology and undersaturation could be distinguished and linked to the different starting conditions. In case the distinction between amorphous precipitate and micro crystals was difficult, polarised light was used (Olympus U-Pot polarisation filter), taking advantage of the birefringence of polarised light caused by the anisotropic structure of the crystals. The crystals then appear bright against a dark background. The boundaries between the precipitation and crystallisation zone and between the crystallisation and the metastable zone were determined by taking the average of the initial protein concentrations of adjacent wells which displayed different precipitation behaviour, i.e. (i) formation of amorphous precipitate or (ii) formation of crystals or (iii) clear solution. After an analysis of the results the experiment was repeated using appropriately adjusted initial protein concentration to reduce the experimental spread at the boundary between two adjacent zones.

5.3.4 Analysis

The concentration of total protein was measured using a Cobas Fara spectrophotometric robot (Roche, Rotkreutz, Switzerland) together with the ESL assay (Roche, Mannheim, Germany), which is based on a reverse biuret method combined with a copper-bathcuproine chelate reaction (Matsushita et al., 1993). The Cobas Fara was programmed to incubate the reagents and samples for 10 minutes at 25°C, then mix them and measure the absorbance at 485 nm after 30 seconds. Protein concentrations were calculated from a bovine serum albumin standard curve. The salts used in the crystallisation studies were shown not to interfere with the assay.

5.4 Results

5.4.1 Phase diagrams in the presence of sodium chloride

The strategy for determining complete phase diagrams was examined at pH 7 in the presence of different concentrations of sodium chloride. This salt was chosen because its impact on solubility is characterised for many proteins (Retailleau et al., 1997), and was considered as a

natural choice to change the ionic strength of a given solution. The time required for the development of the phase transitions observable by microtitre plates was examined first and the progress of the experiment is presented in Figure 5.1. The results in Figure 5.1 A show that in the 3 to 4 hours required for the preparation of the plate (i.e. from filling well A2 to H11), some phase transitions could be observed, particularly at low salt concentrations. At high supersaturations phase transitions occurred more rapidly; particularly the formation of amorphous precipitate was found to happen almost instantaneously (e.g., row A, Figure 5.1 A). After 24 hours, phase transitions were observed for a large number of conditions (Figure 5.1 B). Only wells which contained solutions of low protein and high salt concentration remained clear after 24 hours. There was only a minor further increase in the density of the haze after 96 hours and it was thus concluded that 96 hours was sufficient for the phase transition. After 96 hours, the precipitate formed was analysed by light microscopy and characterised either as being amorphous or crystalline. The different zones so obtained, i.e. precipitation, nucleation and metastable/undersaturated zones are marked in Figure 5.1 C and the corresponding quantitative phase diagram is depicted in Figure 5.2 A in which the solubility data of taken from previous studies (Chapter 2) have been incorporated to complete the diagram. This step is necessary since the microtitre plate method only provides information which initial protein concentration leads to which kind of phase transition and was not used to measure equilibrium concentrations. Although large steps in protein and salt concentration were made from well to well, the data suggests that the metastable zone was relatively wide even after 96 hours of incubation and broadened with salt concentration, meaning that a significant number of the solutions which remained clear were in fact supersaturated. A large nucleation zone was identified at pH 7 which extended throughout the whole range of sodium chloride concentrations examined. The borderline between precipitate formation and nucleation was found to be highly dependent on the presence of salt since amorphous precipitate was only found at very high protein (above 60 mg/mL) and low salt concentrations (below 0.2 M). Above 0.2 M sodium chloride, no amorphous precipitate was formed at the supersaturations tested.

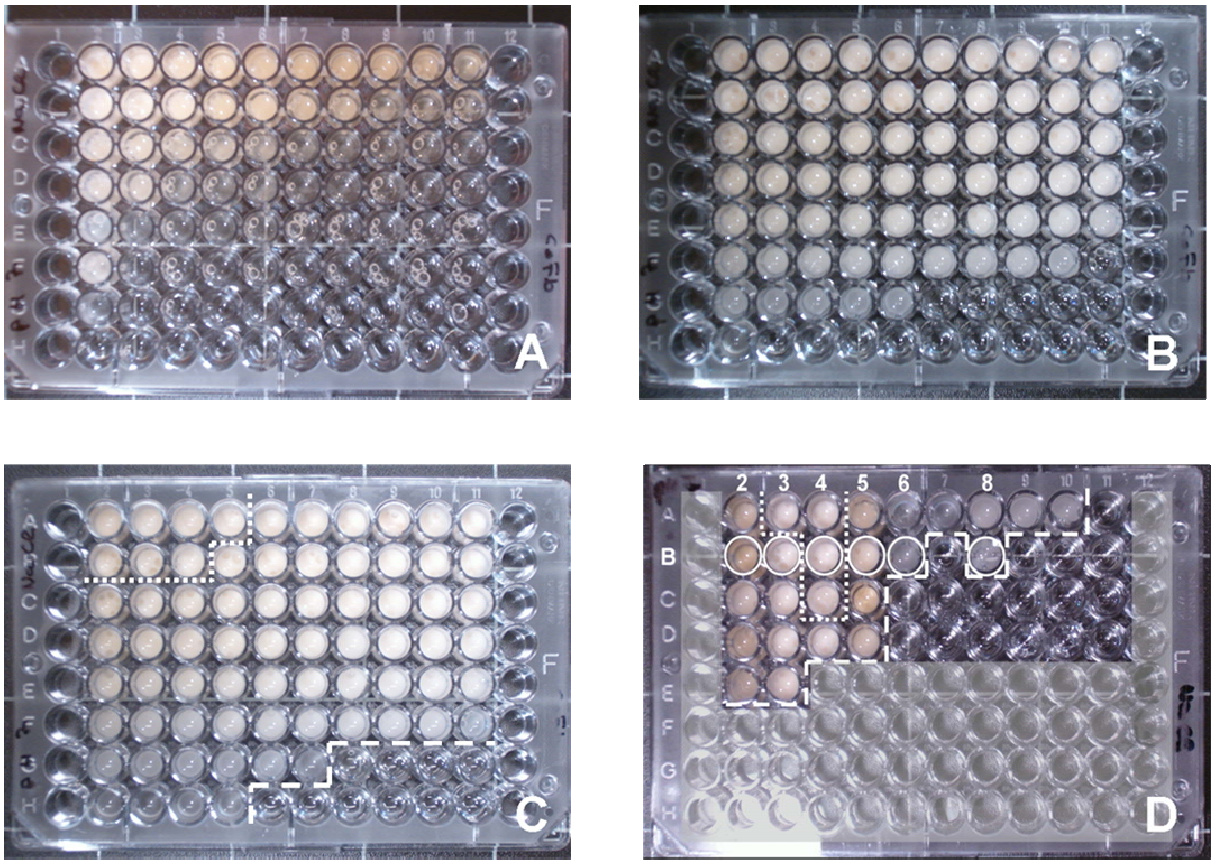

Figure 5.1 Batch crystallisation experiments conducted in a microtitre plate at different times at pH 7 and 40°C. Each well represents a different starting condition. A: at experimental start, B: after 24 hours; C: after 96 hours. On the x-axis the sodium chloride concentration was varied from 0 to 1 M, i.e. 0, 0.05, 0.1, 0.2, 0.3, 0.5, 0.6, 0.7, 0.8 and 1 M, from left to right. On the y-axis the protein concentration was varied within a range of approximately 2 and 80 mg/mL from the bottom to the top in following steps: 2, 4, 10, 20, 30, 40, 60 and 80 mg/mL.

D: second experiment; based on the results of the first experiment, the differences in the initial protein concentration were appropriately reduced to increase the accuracy of the borderline between the regions of different precipitation properties. The dotted lines divide conditions leading to precipitate and crystal formation, the dashed lines separate conditions leading to crystal formation from areas of no phase transition

A very gross discrimination of the different phases was obtained in Figure 5.2 A. Thus, to enhance the accuracy, the experiment was repeated using the same salt concentrations but initial protein concentrations appropriately reduced within the range bars found in the first experiment, keeping the other process parameters constant. The corresponding microtitre plate is shown in Figure 5.1 D and the refined (i.e. based on the second experiment) phase diagram in Figure 5.2 B. Images of the resulting precipitation for conditions at or near the

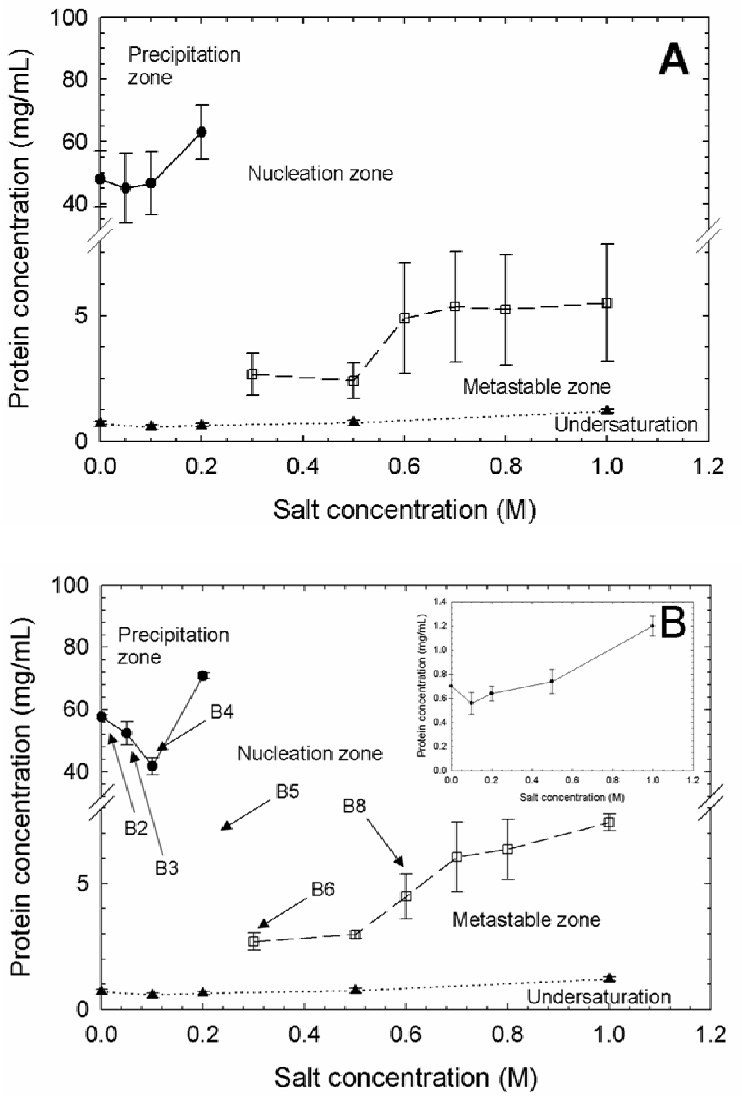

Figure 5.2 **A: Phase diagrams after 96 hours of BHA as a function of sodium chloride concentration at pH 7 and 40°C based on the precipitation properties found in Figure 5.1C. With the exception of the solubility curve, the range bars indicate the difference in initial protein concentration between two adjacent wells of different precipitation properties. (●, solid line) borderline between precipitation and nucleation zone; (□, dashed line): borderline between nucleation and metastable zone; (▲, dotted line) solubility curve which is taken from Chapter 2.**

B: Refined phase diagram of BHA at pH 7 as a function of sodium chloride concentration after 96 hours. The arrows indicate the position of the solutions processed in wells encircled in Figure 5.1 D in terms of salt and protein concentration to demonstrate how the encountered precipitation is incorporated into a phase diagram. With the exception of the solubility curve, the range bars indicate the difference in initial protein concentration between two adjacent wells of different precipitation properties. (●, solid line) borderline between precipitation and nucleation zone; (□, dashed line): borderline between nucleation and metastable zone; (▲, dotted line) solubility curve which is taken from Chapter 2. For the sake of clarity, the solubility curve is redrawn on the upper right-hand corner of Figure 2 B.

borderlines of the different phases are given in Figure 5.3. The distinction between precipitate and crystallisation was found to be difficult in only few cases, e.g., well B3 (Figure 5.3). The typically encountered crystal morphology is rod-like as found in wells B2, B5, B6 and B8 (Figure 5.3). Amorphous precipitate was found in well B4. Only when inspected with great care, could the crystalline structure of the precipitation encountered in well B3 (Figure 5.3) be determined. The edges of the very small crystals were well defined, i.e. they were straight, which is not the case in B4. Furthermore, inspections by polarised light confirmed the crystalline structure of the precipitation found in well B3. In general, the morphology of the crystals was independent of the conditions at which they were grown. However, the rod-shaped crystals differed significantly in size and length-width ratio (Figure 5.3). Very small crystals were formed at conditions close to the precipitation zone, i.e. at high supersaturation (Figure 5.1 D, Figure 5.3, well B3). As the supersaturation was reduced, longer and thinner crystals were observed (Figure 5.3, wells B6 and B8) and the absolute number decreased. When the habit of the crystals formed in well B2 and B3 were compared (Figure 5.1 D, Figure 5.3), a huge difference was found. Crystals of a low length-width ratio and of relatively homogeneous size were grown in well B2 under conditions that were similar to B3. The results indicated that the low salt concentration (no salt in well B2, 0.05 M in well B3) already lead to a large impact on the resulting crystal habit.

It is important to note that only one solid phase was discovered per condition examined. Thus, there was no competition between crystallisation and amorphous precipitation as has been seen e.g., for lysozyme under certain conditions (Feher, 1986). The clear distinction between zones of precipitation and crystallisation was possible in the current work allowing the generation of meaningful phase diagrams.

For the sake of clarity the solubility curve of BHA as a function of sodium chloride concentration at pH 7 is re-drawn as an insert in Figure 5.2 B. The results show that the solubility was very low (between approximately 0.6 and 1.2 mg/mL) throughout the entire salt concentration range studied and passed through a minimum at 0.1 M sodium chloride. The borderline between the nucleation and the metastable zone was clearly distinguishable from the solubility curve and followed a similar trend, but was seen to broaden with increasing salt concentration (Figure 5.2 B). The borderline between nucleation and metastable zone at salt concentrations lower than 0.3 M could not be determined due to the

presence of crystals. Thus, no information can be provided on whether the metastable zone also passed through a minimum but is expected to be located at a protein concentration between 1 and 3 mg/mL. Interestingly, the borderline between precipitation and the nucleation zone also passed through a minimum (relative supersaturation of 74) at 0.1 M sodium chloride (Figure 5.2 B). At salt concentrations higher than 0.2 M, no amorphous precipitation but crystals were formed under the tested conditions.

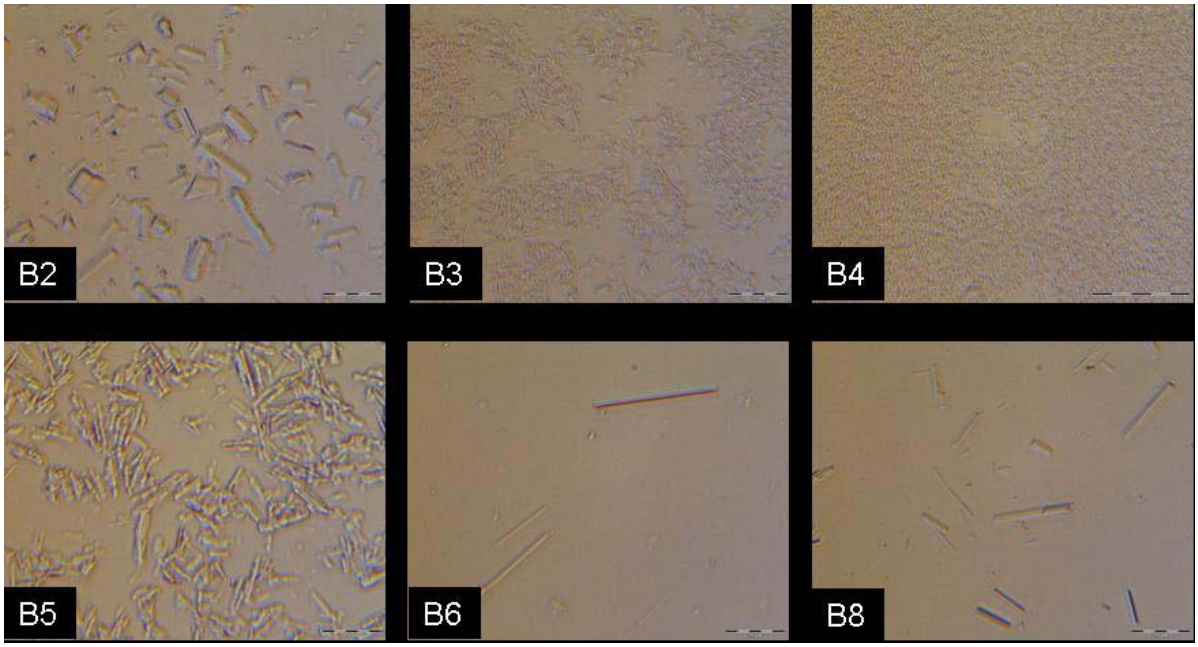

Figure 5.3 Light microscopy analysis at a 400-fold magnification of BHA precipitation resulting in wells B2 (0 M), B3 (0.05 M), B4 (0.1 M), B5 (0.2 M), B6 (0.3 M) and B8 (0.6 M) of the microtitre plate depicted in Figure 5.1 D (corresponding wells encircled). The position of these solutions is given in the phase diagram shown in Figure 5.2 B. The scale bars represent a length of 20 µm.

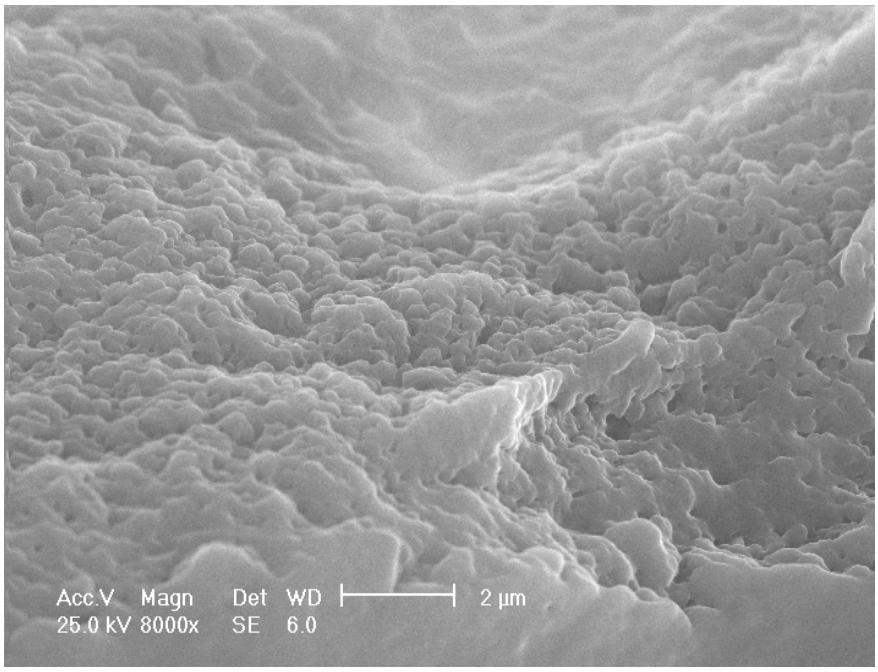

Figure 5.4 Amorphous precipitation of BHA obtained from an initial protein concentration of 60 mg/mL, no salt, pH 6 at 40°C, picture taken by environmental scanning electron microscopy (ESEM).

5.4.2 Sodium thiocyanate dependence of the phase diagram at pH 7

Thiocyanate is well known to be a poor precipitant and its presence often leads to large increases in protein solubility under certain conditions (Chapter 2). It was thus of interest to examine its effect on the precipitation, nucleation and metastable zone. At pH 7, the BHA solubility did not pass through a minimum but increased approximately 20-fold as the sodium thiocyanate concentration was raised to 1 M (Figure 5.5 A). The metastable zone was small at 0.1 and 0.2 M sodium thiocyanate but broadened significantly with increasing salt concentration (Figure 5.5 A).

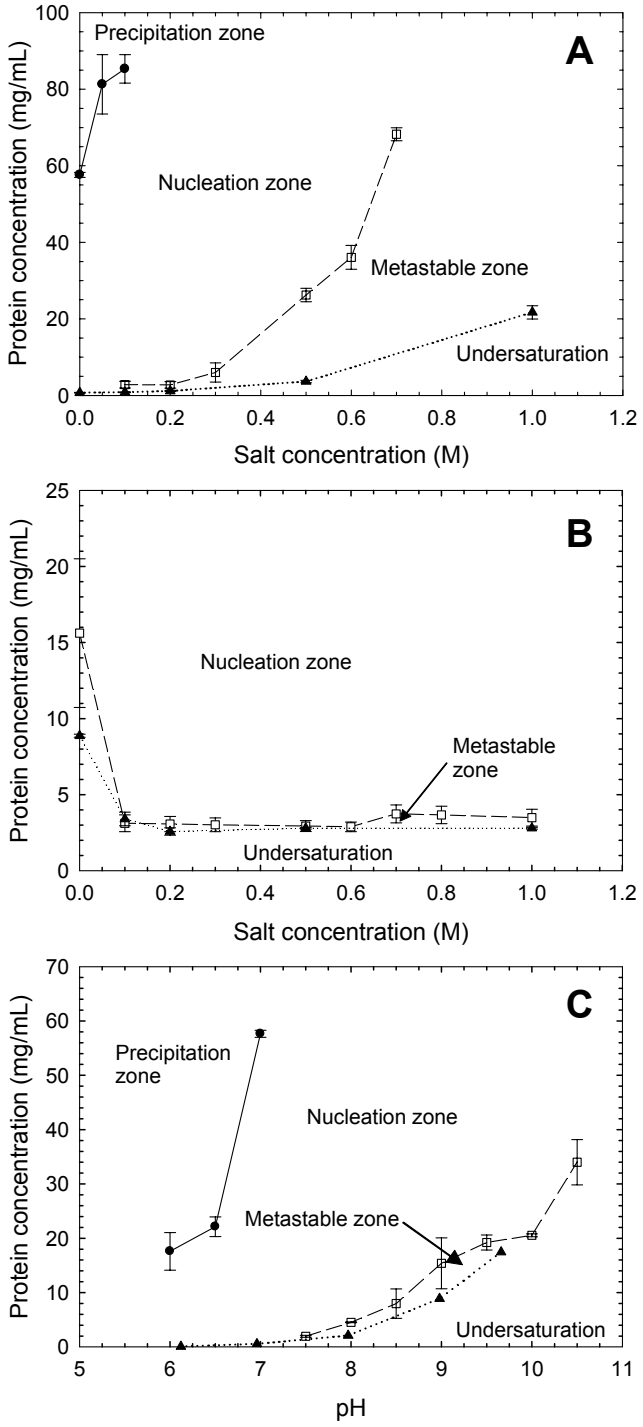

Figure 5.5 **Refined phase diagrams of BHA after 96 hours at 40°C as a function of A: sodium thiocyanate at pH 7; B: sodium chloride at pH 9; C: pH without added salt. (●, solid line): borderline between precipitation and nucleation zone; (□, dashed line): borderline between nucleation and metastable zone; (▲, dotted line): solubility curve. With the exception of the solubility curves which were taken from Chapter 2, the range bars indicate the difference in protein concentration between two adjacent wells of different precipitation properties, i.e. amorphous precipitate, crystals or no phase transition.**

A broad nucleation zone was observed throughout the entire salt concentration range examined. As was observed with sodium chloride addition at the same pH (Figure 5.2 B), amorphous precipitate only formed at very low salt (between 0 and 0.1 M) and high initial protein concentrations. The borderline between the precipitation and nucleation zone did not pass through a minimum but rapidly increased with salt concentration (Figure 5.5 A). The relative supersaturation needed to induce the formation of amorphous precipitate ranged from 82 for the thiocyanate-free solution to 101 at 0.1 M. The crystal habit was rod-shaped but a more complex size distribution was encountered than when sodium chloride was used. In general, the crystals became bigger with increasing salt concentration. However, the crystal size was less homogeneous compared to when sodium chloride at pH 7. With decreasing supersaturations, some long BHA-crystals were seen (approximately 15 μm) but small ones of approximately 5 μm in length were also found. The number of crystals decreased with decreasing supersaturation.

5.4.3 Phase diagram at pH 9 as a function of sodium chloride concentration

The influence of pH on the phase diagram of BHA in the presence of sodium chloride was examined at pH 9 under conditions otherwise identical to those employed at pH 7. The results in Figure 5.5 B show a very large difference in the phase diagram at pH 9 than was seen at pH 7 (Figure 5.2B). The use of the higher pH drastically reduced the width of the metastable zone seen at pH 7. The addition of 0.1 M sodium chloride at pH 9 resulted in a sharp decrease of the solubility to approximately one third of the original value, further salt addition had almost no effect on solubility (Figure 5.5B). The boundary between the nucleation and metastable zone was located very close to the solubility curve. The calculated relative supersaturations needed to induce nucleation at 0.3 and 0.5 M were very small, i.e. 1.05 in both cases and only slightly increased to values of approximately 1.3 at 0.7 and 1 M. In contrast, the minimum relative supersaturation required to induce nucleation at pH 7 ranged between 4.1 and 6.2. The crystal habit and morphology did not change upon salt addition. The commonly observed rod-shaped crystals of similar size were also found here and are the same as seen in Figure 5.3.

5.4.4 Effects of pH on the phase diagram

A phase diagram for BHA as a function of pH (no salts added except the buffer) was generated and is presented in Figure 5.5 C. The solubility increased by a factor of approximately 200 when changing pH from 6 to 10 (Figure 5.5 C). The metastable zone was found to be small and paralleled the solubility curve. The relative supersaturation needed for crystal formation ranged between 1.3 at pH 9.5 and 1.9 at pH 8. A large nucleation zone for BHA was found and crystals were formed under a large combination of conditions. High relative supersaturations, i.e. between 67 at pH 6 and 82 at pH 7 were necessary to provoke the formation of amorphous precipitate. The highest initial protein concentration tested (approximately 80 mg/mL) was not enough to induce the formation of amorphous precipitate at pH values above 7. However, it is to be accepted that the strong dependence on pH of the precipitation zone towards higher pH-values would continue. The observations showed that the crystals became smaller in size when the edge of the precipitation zone was approached and longer and fewer in number the closer to the metastable zone they were grown. The crystal morphology did not change as a function of pH.

5.5 Discussion

5.5.1 Precipitant effects on the phase diagram

The microtitre plate method employed here together with traditional batch crystallisation methods conducted in 1.5 mL Eppendorf tubes has permitted the generation of complete phase diagrams. At pH 7 with increasing sodium chloride concentration the BHA solubility slightly deceased from 0.7 to 0.8 mg/mL from zero to 0.1 M but increased to 1.2 mg/mL at 1 M. In the same concentration range and same pH, the solubility increased from 0.7 mg/mL to 22 mg/mL when sodium thiocyanate was used instead. At pH 9, the solubility decreased from 15 to 3 mg/mL upon addition of 0.1 M sodium chloride and a further increase in salt concentration to 1 M did not cause any notable changes. When no salt except the buffer was present in the solution, the solubility increased from 0.25 mg/mL at pH 6 to 15 mg/mL at pH 9.5. The reasons for the solubility behaviour observed have been examined in earlier studies (Chapter 2).

The true value of the microtitre plate method appears to be the ability to rapidly observe and measure the complete phase diagram which we show for the first time. In this current work different conditions lead to major changes in the metastable zone. Increasing sodium chloride concentrations broadened the metastable zone at pH 7 (Figure 5.2 B), but increasing the pH to 9 significantly reduced the width of the metastable zone in the presence of salt (Figure 5.5 B) but had little effect in the absence of salt (Figure 5.5 C). Shifting to the very poor precipitant thiocyanate greatly broadened the metastable zone (Figure 5.5 A).

The metastable zone is defined as an area in the supersaturated region in which the crystallisation solution remains supersaturated without the formation of a solid phase for an infinite length of time if it is left undisturbed (Chayen, 2005) but can produce a shower of fine crystals e.g., when the solution is mechanically shocked (Feigelson, 1988). To form stable nuclei, a certain energy barrier has to be overcome. The energy is provided by the difference in chemical potential which is determined by the degree of supersaturation (Arakawa & Timasheff, 1985). If the supersaturation is very low, the energy barrier may be too high so that stable nuclei will only form after an infinite time. The reason for that is that the size of the critical nuclei gets larger with decreasing supersaturation at which the energy provided by the difference in chemical potential is not high enough to stabilise smaller nuclei (McPherson et al., 1995). At lower supersaturations, the probability to form larger nuclei is reduced since more molecules have to be correctly incorporated into the crystal array (Malkin et al., 1996). At the same time, the amount of free protein molecules available to form the critical nuclei is smaller at lower supersaturations. Both effects decelerate the formation of the critical nuclei so that crystals may only appear after a long period which exceeds the experimental run-time.

The experimental run-time used in this study was 96 hours and the crystallisation solutions were continuously shaken so that a certain amount of energy was constantly introduced into the system. The definition of the metastable zone employed in the current study may therefore deviate from the one given above but is justified by the aim to generate process-relevant data. Since the free energy needed to form stable nuclei is the decisive unit to determine the width of the metastable zone (Arakawa & Timasheff, 1985), possible influences on this energy caused by the addition of salts are illuminated further here. Phenomena governing the solubility are also believed to influence the width of the metastable zone. BHA has a pI of 6 which means that the net charge at pH 7 and 9 is negative but the magnitude is bigger at pH 9

(Chapter 2) which is also reflected by the big difference in solubility (0.7 mg/mL at pH 7 and 9 mg/mL at pH 9). At pH 7, the addition of sodium chloride only leads to minor effects on solubility by charge screening by sodium ions due to the lower charge. At this pH, increasing repulsive forces above 0.1 M is believed to be caused by the accumulation of sodium ions in the vicinity of the protein molecule as also seen for apoferritin (Petsev & Vekilov, 2000), characterised at approximately one pH-unit above the pI, similar to the conditions currently discussed. The contribution from the chloride ions is small because of its position in the middle of the Hofmeister series, meaning that it is neither kosmotropic nor chaotropic. The increase in solubility may appear small in absolute terms but doubles from 0.1 to 1 M which is a clear indication that the repulsive forces are increasing with raising sodium chloride concentrations. The increasing repulsive forces are also responsible for the broadening of the metastable zone (Figure 5.2 B). The minimum relative supersaturations required to induce nucleation increase from 4.2 at 0.3 M to 6.2 at 1 M sodium chloride concentration. Thus it appears that augmenting repulsive forces caused by increasing sodium chloride concentrations are not only provoking an increase in solubility but that the presence of sodium chloride increases the amount of free energy needed to form the critical nuclei which results in a wider metastable zone.

When thiocyanate is used instead, both the solubility and the width of the metastable zone dramatically increase with raising salt concentration. Thiocyanate is a strong chaotrop and influences the solubility by binding to oppositely charged amino acid residues. Thus, it adds more negative charges to the protein surface which leads to strongly increased repulsive forces. Concurrently, both the solubility and the width of the metastable zone are increasing. In the presence of thiocyanate, effects caused by sodium as found for sodium chloride at this pH are negligible. The strongly repulsive forces have a negative effect on the formation of the critical nuclei which leads to a higher energy demand which can only be provided at higher supersaturations.

At pH 9 the addition of 0.1 M sodium chloride leads to a strong decrease in solubility which is caused by charge screening of negatively charged patches on the surface of the molecule by sodium ions. In contrast to pH 7, further adding sodium chloride does not lead to an increase in solubility. Although the absolute solubility is higher at pH 9 (approximately 3 mg/mL compared to 0.6 mg/mL at pH 7, both in the presence of 0.1 M sodium chloride), the creation of repulsive forces e.g., by the mechanisms presented above leading to an enhanced solubility

seems to be more difficult at pH 9 than at pH 7 and probably requires salt concentrations higher than 1 M. Such a behaviour has been seen for lysozyme the solubility of which only started to increase at sodium chloride concentrations far above 1 M after having shown a pronounced salting-out (Bénas et al., 2002). Accordingly, the metastable zone is much smaller and less dependent on salt concentration at pH 9 due to the favoured nuclei formation compared to pH 7 at least within the examined salt concentration range. In light of this, it is noteworthy that the boundary between nucleation and metastable zone above sodium chloride concentrations of 0.5 M is located at higher absolute protein concentrations at pH 7 compared to pH 9.

The very small metastable zone found at pH 9 within sodium chloride concentrations of between 0 and 1 M would be in contrast to the general assumption that protein crystal growth requires a much higher supersaturation than small molecules. Supersaturation ratios of 1.1 are sufficient to induce the growth of potassium sulphate solutions (Sha et al., 1996) whereas the minimum supersaturations necessary to induce the formation of protein crystals are typically one order magnitude higher (Chernov, 1997; Feigelson, 1988). However, thaumatin crystals were obtained at supersaturation ratios between 3.0 and 5.5 (Malkin et al., 1996) and lysozyme crystals were grown at relative supersaturations as small as 1.8 (Ataka & Tanaka, 1986). The low relative supersaturations, necessary to crystallise BHA under the conditions employed here, in particular in the presence of sodium chloride at pH 9, (supersaturation ratio of 1.05) may be a result of heterogeneous nucleation e.g., on dust particles or other impurities or secondary nucleation e.g., by the continuous power input through stirring and consequent collision between of two crystals or between crystals and the well-surface. Heterogeneous and secondary nucleation can to a large extend be ruled out for vapour-diffusion methods as hanging- and sitting-drops since the systems employed have to be completely closed but are believed to be the by far dominating processes in stirred open batch crystallisations (Tavare, 1986). The lysozyme crystals obtained at an initial supersaturation ratio of 1.8 were also grown in batch mode (Ataka & Tanaka, 1986), albeit unstirred. In batch crystallisations, the supersaturation is induced by increasing the precipitant concentration to move the system across the solubility curve. In contrast to hanging or sitting drop, it is practically unavoidable to create concentration gradients upon precipitant addition such that high local supersaturations are created and stable nuclei formed. This fact may result in a situation in which otherwise identical crystallisation experiments are successful in batch mode but fail in hanging or sitting drop experiments because the solution was in fact either undersaturated or

metastable. This would lead to the not quite correct conclusion that batch crystallisations require lower supersaturations than hanging or sitting drop crystallisations.

In the absence of salts, the width of the metastable zone was essentially unchanged with increasing pH (Figure 5.5 C). Changing the pH e.g., to alkaline conditions provoked the continuous deprotonation of charged amino acid side groups, thereby shifting the protein net charge to more negative values. The repulsion between the protein molecules was increased which consequently resulted in an enhanced solubility. Although ions have to be added to change the pH, i.e. the counter ions to protons or hydroxides, their concentration is by far lower as is the case for salt addition. The mechanisms as discussed above by which the different ions change the solubility and the width of the metastable zone are not believed to be relevant in this case or only to a negligible degree. It seems that increasing repulsive interactions as generated by increasing pH do not necessarily lead to a wider metastable zone but the presence of ions plays a decisive role in the generation of the metastable zone beyond the induction of attractive or repulsive interactions, arguably by modifying the amino acid residues involved in forming the contact interfaces between protein molecules in the crystal lattice.

The results of this study demonstrate that the borderline between nucleation and precipitation zone almost parallels the solubility curve (Figure 5.2 B and Figure 5.5 A and C). This leads to the conclusion that the mechanisms governing the solubility are also involved in the precipitate formation. In contrast to crystals, amorphous precipitate can be approximated as one-dimensional chains of randomly orientated molecules and thus, the energy required for growth is size-independent (Feher & Kam, 1985). The growth is induced by strong non-specific van der Waals forces which increase with the degree of supersaturation. Beyond a certain supersaturation, they become too strong such that the molecules cannot form ordered arrays leading to crystals but amorphous precipitate is formed. The van der Waals forces are moderated by repulsive forces which are induced e.g., by higher net charges and/or the presence of salts causing higher solubilities. In light of this, it is understandable that the borderline between the precipitation and the nucleation zone follows similar trends as the solubility curve.

The relative supersaturations required for the formation of amorphous precipitate (minimum relative supersaturation of 67 measured at pH 6, no salt) are considerably higher than e.g., for

lysozyme, which amorphously precipitates already at relative supersaturations of 10 at pH 4.2 and approximately 1.4 M sodium chloride concentration. However, with decreasing salt concentrations, the relative supersaturations necessary to induce amorphous precipitation get significantly enhanced and reach values comparable with BHA (Feher & Kam, 1985).

A recently developed model describes the phase behaviour of globular proteins using the second order perturbation theory of Henderson and Barker which assumes proteins carrying asymmetric surface charge distributions in the presence of a monovalent electrolyte (Tavares et al., 2004). The model focuses on electrostatic effects consisting of charge-charge repulsion, and charge-dipole and dipole-dipole attractions. Although the model aims to predict liquid-liquid phase separation, a phenomenon which has not been observed for BHA, it is assumed that the calculated liquid-liquid phase separation lines and the borderline between the nucleation and the precipitation zone determined in this study to a large extent describe the same phenomenon, as has previously been suggested for the enzyme rennin (Bunn et al., 1971). According to the model, electrostatic attractions can be maximised at salt concentrations typically in the range of between 0.01 and 0.1 M. Furthermore, the model suggests that if the net charge on a protein was increased (e.g., by a pH change) it would lead to a less pronounced maximum in attractive electrostatic interactions which would eventually vanish if the protein had a very high net charge. Both suggestions would be in very good agreement with the findings of this study. A minimum both in solubility and the borderline between amorphous precipitation and nucleation zone was observed for BHA at 0.1 M sodium chloride at pH 7. At higher pH, no amorphous precipitation was observed which would be in line with the model's prediction that increased net charge reduces the maximum in attractive electrostatic interactions. At pH 7, the addition of sodium thiocyanate did not yield a minimum in solubility nor lead to a reduction in the borderline between precipitation and the nucleation zone. The ability of thiocyanate to bind to the BHA molecule would increase the net charge with exactly the consequences as suggested by the model. The current study as well as the model described above emphasise the important role of repulsive and attractive interactions of charged globular proteins which strongly depend on pH and salt concentration.

5.5.2 Crystal habit

Another beneficial feature of the microtitre plates is that the examination of the precipitate or crystals formed can also be used to characterise of the crystal morphology, where applicable. It was found that the crystal morphology was independent of the growth conditions since only rod-like crystals were formed. Crystals only differed in size, distribution and length-width ratio. The crystal size was mainly dominated by the supersaturation. Initial supersaturations close to the nucleation zone led to the formation of small crystals whereas very long and homogeneous crystals were found if grown on supersaturations near the metastable zone (Figure 5.3). This trend has also been found for lysozyme (Ataka & Tanaka, 1986), and is in agreement with the commonly accepted considerations of crystal growth (Chernov, 1997; Malkin et al., 1996; McPherson et al., 1995). At higher supersaturations the size distribution was commonly broader but grew more homogeneously at lower supersaturations. When the crystallisation process is started from higher supersaturations, the nucleation rate is also high in the beginning so that many crystals are formed the growth of which may cease sooner as lattice errors are more likely to occur at higher supersaturations. Crystals formed at later stages of the process start growing from lower supersaturations leading to larger sizes (Weber, 1991). As a result, the size distribution broadens. The addition of sodium thiocyanate led to very broad size distributions, not only at high initial supersaturations. The mechanism described above is also relevant here but to an even larger degree. The binding of thiocyanate ions to the protein surface and the resulting repulsive interactions lead to a retarded formation of critical nuclei and crystal growth. The slow degradation of the supersaturation causes the formation of crystals very different in age and the growth would be initiated at different supersaturations leading to different crystal sizes at the termination of the experiment (Boistelle & Astier, 1988). For thiocyanate this would also be true for crystallisation processes started from lower initial supersaturations. Another observation was that the addition of sodium chloride caused the growth of very thin crystals whereas more compact crystals were formed without salt and with thiocyanate. It is suggested that the different salts added impose specific steric hindrances which depend on the amino acid residues constituting the contact between the molecules within the crystal lattice such that either the growth in the longitudinal direction gets favoured or transverse growth is decelerated upon sodium chloride addition.

5.5.3 Relevance and further areas of application of the microtitre plate method

Microtitre plates provide important information on kinetic aspects of the phase separation process. From a process point of view, it is of high interest on how long a crystallisation process should be conducted. It was found that amorphous precipitate appeared instantaneously after the supersaturation was induced, a property commonly observed for proteins (Feher & Kam, 1985). When starting a large-scale crystallisation process, tnstantaneously increased turbidity may indicate the formation of amorphous precipitate and could be used as a criterion to stop the process immediately. At conditions under which amorphous precipitate is once formed, crystal formation is unlikely throughout the processing time (Ng et al., 1996). Moreover, in case of BHA, a batch crystallisation process would, for a majority of conditions, be close to be finalised after 24 hours; an extended processing time may not lead to significantly increased yields. In addition, the microtitre plate method also provides information how to slow down the crystallisation process and thus widen the metastable zone. Crystallisation processes of very slow kinetics would offer the possibility of concentrating the protein beyond the solubility limit with a calculable risk of phase separation within a certain time. The addition of appropriate amounts of selected salts may thus be one way to steer the width of the metastable zone which would enhance the possibilities to generate high concentration formulations opening up new application areas of proteins in industry.

Next to the generation of entire phase diagrams, the approach presented in the current study allows screening of many growth conditions which may eventually lead to the formation of crystals of desired properties. Organic additives like polyethylene-glycol (PEG) (Tardieu et al., 2002) of various molecular weights or 2-methyl-2-4-pentanediol (MPD) (Anand et al., 2002) are compounds often present in crystallisation screening kits and may also be beneficial in large-scale crystallisation. Naturally, the additives found to lead to the desired crystal size and distributions have to be carefully inspected for possible disadvantages further down in the downstream process. However, if e.g., the influence of the two compounds mentioned above on the crystal quality is to be tested together with different other salts at various pH-values, the number of necessary experiments rapidly increases to an unjustifiable degree. In such a case, experimental design based on factorial approaches (Betts et al., 1989; Carter, Jr. & Carter, 1979) should be used. Recent developments of neural networks have shown that they can be useful for analysing complex relationships between a large number of variables with

respect to the results these combinations produce (DeLucas et al., 2005). Although directed towards predicting optimum conditions likely to yield to diffraction quality crystals, neural networks could also be trained to indicate crystallisation conditions leading to large crystals of narrow size distributions. Microtitre plates provide an excellent platform for computer-aided experimental design coupled with automated liquid handling systems which would increase the experimental accuracy, reduce the time and labour demand and would offer the possibility to further reduce the scale, hereby reducing the protein consumption. However, the approach presented here relies on the manual inspection of the resulting precipitate, i.e. to distinguish between crystals and amorphous precipitate. The use of advanced picture recognition software may be a promising way completely automating the procedure. The generation of complete phase diagrams requires the knowledge of solubility curves which have to be generated separately, e.g., by traditional batch crystallisations. The use of microtitre plates provides information on the borderline between amorphous precipitation and nucleation zones and nucleation and metastable/undersaturated zones. The phase diagrams so obtained are thus not complete and do not provide information about the width of the metastable zone and the position of the solubility curve. However, the resulting phase diagrams are still of high value for the design of crystallisation processes as they provide information on how to avoid the formation of amorphous precipitation, which crystal habit and morphology to expect and whether the process can be finalised within the given time frame.

The relevance of the results obtained in micro-scale has to be put into perspective with regard to the applicability in large scale crystallisations. The scale-up of protein crystallisation processes has not been the subject of many published studies. The scale-up of a crystallisation processes of an alcohol dehydrogenase and a lipase from 1 mL to 500 mL has been described as simple and straightforward (Lee et al., 2000). In contrast, the scale-up to litre or even cubic meter crystallisation processes has been found challenging and required systematic approaches for success (Basu et al., 2004; Schmidt et al., 2005). It is believed that the phase diagrams generated in microtitre plates are well suited as a guide to large scale crystallisation processes. It still needs to be verified in how far the phase diagrams generated in small scale correlate with the phase behaviour in large scale. Only if they correlate reasonably well, scale-up scale-up strategies can be based on information provided by experiments in micro-scale.

5.6 Conclusions

Microtitre plates together with conventional batch crystallisations to determine solubility curves have been shown to be time and material saving tools to generate complete phase diagrams consisting of precipitation, nucleation, metastable and undersaturated zones. The analysis of the resulting phase diagrams revealed that BHA crystallised under a large variety of conditions. Amorphous precipitate was formed only at very high relative supersaturations. The borderlines between the nucleation and precipitation zone followed similar trends as the solubility curves. The width of the metastable zone could be manipulated by salt addition. The interplay between the precipitation, nucleation and metastable zones along with the solubility curve as determined by this study suggests that the phase behaviour of BHA is mainly governed by electrostatic interactions. The findings of this study are in good agreement with the predictions of a mechanistic model which focuses on the description of electrostatic interactions assuming globular proteins of asymmetric surface charge distributions. The morphology of the crystals was independent of the growth conditions tested. The supersaturation was shown to be the decisive factor governing the size and size distribution of the crystals formed, although the tested salts had their individual effects on the crystal habit. From a processing point of view, a powerful approach is presented in this study for estimating the formation of amorphous precipitate, where to find favourable crystallisation conditions, which crystal habit to expect, how to manipulate the kinetics of the crystallisation process or how to circumvent any kind of phase transition. The method presented in the current study can be further scaled-down by conducting the experiments at even smaller working volumes (i.e. < 160 μL) to minimise the material consumption and by coupling with automated pipetting systems the data accuracy can be enhanced and the time demand for plate preparation further reduced. Moreover, microtitre plates provide an excellent platform to conduct experiments based on factorial design which enables the simultaneous examination of complex interactive relationships between a large number of variables with respect to their impact on the phase behaviour yet at a limited number of trials. Although the scale-up to pilot-plant or production volumes was not the subject of this study and the correlation between the phase behaviour in μL and large scale still needs to be verified, the methods developed in the current study are considered to have the potential to become indispensable tools to develop successful scale-up strategies for the rapid and rational design of truly optimised protein downstream processes.

6 The potential of self-interaction chromatography for the rapid determination of the second osmotic virial coefficient in aqueous preparations of *Bacillus halmapalus* α-amylase

Cornelius Faber,[a,b] André Dumetz,[e] Timothy J. Hobley,[b,*] Jørgen Mollerup,[c] Owen R. T. Thomas,[b,d] Svend G. Kaasgaard,[a] and Abraham M. Lenhoff[e]

[a] Novozymes A/S, Novo Alle, DK-2880 Bagsværd, Denmark

[b] Center for Microbial Biotechnology, BioCentrum-DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[c] Department of Chemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[d] Department of Chemical Engineering, School of Engineering, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

[e] Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716, USA

*Corresponding author

e-mail: th@biocentrum.dtu.dk (T. Hobley)

Telephone: +45 45 25 27 06

Fax: +45 45 88 41 48

Keywords: amylase, crystallisation; immobilisation; protein; second osmotic virial coefficient; self-interaction chromatography;

6.1 Abstract

Self-interaction chromatography (SIC) has recently been proposed as a highly efficient method for measuring the second osmotic virial coefficient B_{22} of proteins in aqueous solutions. The applicability of this method for a semipure preparation of *Bacillus halmapalus* α-amylase (BHA) has been investigated for a variety of different solution conditions. Light scattering techniques, which are typically employed to determine B_{22}, encounter problems if the protein solution under investigation has strong tendencies to form soluble aggregates or is of limited purity. In this regard, SIC was shown to be unaffected and led to reproducible results. The B_{22} was determined in the presence of selected anions and cations from the Hofmeister series at conditions up to 1 M. The good crystallisability of BHA observed in solubility experiments was consistent with SIC results, which for the majority of conditions suggested B_{22} values within the crystallisation slot. Trends predicted by SIC were in qualitative agreement with classical measurements of solubility, S. The Hofmeister series for anions and cations were correctly reflected by B_{22}. Major discrepancies between solubility and SIC measurements were seen only for lithium and azide. Despite regions of significant scatter, the experimentally found correlation between B_{22} and S could sufficiently well be described by the Haas-Drenth-Wilson model. The established correlation between B_{22} and S is believed to be strong enough to allow solubility to be replaced by B_{22} measurements to take advantage of the distinctive properties of SIC, such as the significantly reduced protein and time demand. This study emphasises the large potential of SIC, e.g., in screening routines for desired solution properties of novel proteins or as a tool in process control in all stages of the production of bulk proteins.

6.2 Introduction

The advent of novel methods and techniques to rapidly develop and produce proteins tailored for specific applications, e.g., in the pharmaceutical or food industry, has increased the challenges faced in downstream processing. In many cases, highly concentrated protein formulations are desirable in order to maximise the specific volumetric activity which, e.g., lowers expenses in transportation and storage and opens up new fields of applications (Basu et al., 2004). At the same time, the efficient use of production facilities requires unit

operations to be run at or near the solubility limit, thereby increasing the risk of uncontrolled crystallisation or precipitation, which could lead to product loss or delays in production and delivery. The situation is exacerbated by the availability neither of solubility data nor tools to measure solubility within a short time to take control of production processes. As a consequence, there is the strong desire for alternative methods to rapidly gather information on the phase separation properties of any given protein solutions, be it in production processes to identify and avoid critical operation conditions or in early screening routines to discover promising target molecules of desired properties to rationalise the design of the downstream processing.

Instead of focusing on the solubility, a possible alternative would be to measure the second osmotic virial coefficient, B_{22}. This quantity represents a Boltzmann-weighted average measure of protein-protein interactions; positive B_{22}-values correspond to repulsive and negative to attractive interactions (Tanford, 1961). In light of this, there is experimental evidence that protein solutions characterised by negative B_{22} values lead to phase separations whereas the opposite is true for positive B_{22} values. Protein solutions that give rise to slightly negative B_{22} values, i.e. from -0.5 to -8×10^{-4} mol mL g^{-2}, were shown to have an enhanced propensity to crystallise; consequently, this range was termed the "crystallisation slot" (George & Wilson, 1994). A general correlation between B_{22} and solubility for several globular proteins has been demonstrated (Guo et al., 1999), which has been described in quantitative terms (Haas et al., 1999; Ruppert et al., 2001). The B_{22} value would thus be a convenient quantity to use in process control to operate fermentation and downstream processes within a range of desired conditions, e.g., to either avoid separation processes or to guide crystallisation.

Methods to measure B_{22} are, e.g., small-angle X-ray (Ducruix et al., 1996) or neutron scattering (Velev et al., 1998), membrane osmometry (Tessier & Lenhoff, 2003; Vilker et al., 1981) and sedimentation equilibrium (Jelesarov et al., 1998). These methods are generally labour-intensive and expensive in terms of both time and protein and thus unsuitable to be employed in process control. Probably the most widespread method for the determination of B_{22} is static light scattering (SLS) (Georgalis & Saenger, 1999; Wyatt, 1993). When implemented in a higher throughput flow technique (Wilson, 2003) and integrated into an automated measuring system, this method may have the potential to be used for on-line process control. In spite of its technical sophistication, however, SLS suffers from a principal

problem: since the scattering intensity varies strongly with the diameter ($I \sim d^6$, Rayleigh's law) of the scatterer, B_{22} measurements in polydisperse systems, which are frequently encountered in bulk protein production processes, are problematic. In particular, soluble aggregates generate signals that may lead to irrelevant B_{22} values.

This problem could be circumvented or at least minimised by employing a method recently developed, namely self-interaction chromatography (SIC). This method involves covalently immobilising protein molecules on chromatographic particles, packing these particles into a chromatographic column, and passing the same protein through the column (Patro & Przybycien, 1996). The retention time reflects the average protein interactions between free and immobilised protein and is linked to B_{22} through the potential of mean force. B_{22}-measurements by SIC have been shown to agree quantitatively with those made by SLS and membrane osmometry but at least one order of magnitude more efficient in terms of protein consumption and time (Tessier & Lenhoff, 2003). Miniaturisation to lab-on-chip scale has been shown to be feasible which would lead to a further reduction in protein consumption of up to two orders of magnitude. In this study, the applicability of SIC to measure the B_{22} of semipure aqueous solutions of recombinant *Bacillus halmapalus* α-amylase (BHA) and the feasibility to replace laborious solubility measurements by SIC were evaluated. This model protein was chosen because its crystallisation and solubility properties are well characterised under many conditions (Chapter 2) so that trends in B_{22} and solubility can be compared and correlations established.

This problem could be circumvented or at least minimised by employing a method recently developed, namely self-interaction chromatography (SIC). This method involves covalently immobilising protein molecules on chromatographic particles, packing these particles into a chromatographic column, and passing the same protein through the column (Patro & Przybycien, 1996). The retention time reflects the average protein interactions between free and immobilised protein and is linked to B_{22} through the potential of mean force. B_{22} measurements by SIC have been shown to agree quantitatively with those made by SLS and membrane osmometry but they are at least an order of magnitude more efficient in terms of protein consumption and time (Tessier & Lenhoff, 2003). In this study, the applicability of SIC to measure the B_{22} of semipure aqueous solutions of recombinant *Bacillus halmapalus* α-amylase (BHA) and the feasibility to replace laborious solubility measurements by SIC were evaluated. This model protein was chosen because its crystallisation and solubility properties

are well characterised under many conditions (Chapter 2) so that trends in B_{22} and solubility can be compared and correlations established.

6.3 Materials and methods

Recombinant *Bacillus halmapalus* α-amylase (BHA) was expressed in *Bacillus licheniformis* grown on a complex medium, and purified at Novozymes A/S, Bagsværd, Denmark. Cells and other solids were removed by filtration; the resulting solution was concentrated by ultrafiltration (10 kDa cut-off), and subsequently crystallised by lowering the pH to 7.5. The crystals were then collected by centrifugation. A small fraction of the material was further purified to homogeneity, referred to as "purified BHA". The remaining part was not subjected to any operations for further purification, and denominated as "crude BHA".

6.3.1 Reagents

Acetic acid, sodium hydroxide pellets, EDC (1-ethyl-3-(dimethylamino) propyl carbodiimide), HEPES (4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid), lithium chloride, MES (2-(N-morpholino)ethanesulfonic acid), NHS (N-hydroxysuccinimide), potassium chloride, sodium nitrate, sodium sulphate and sodium thiocyanate were obtained from Sigma (St. Louis, MO, USA). Boric acid and sodium chloride were purchased from Fisher Scientific (Pittsburgh, PA, USA). All chemicals were of analytical grade. Proxel LV was from Avecia, Manchester, UK.

6.3.2 Polydispersity

The polydispersity of BHA solutions of interest were measured in a Zetasizer Nano ZS (Malvern Instruments, Malvern, UK). Samples were, when necessary, diluted 10-fold in the selected buffer and filtered through a 0.22 μm PTFE membrane (Millipore Corp., Bedford, MA, USA). The buffers used for dilutions were of the same composition as the samples with respect to pH, salt type and concentration. 1.5 mL of the diluted samples were transferred into

microcuvettes (Sarstedt, Nümbrecht, Germany) previously cleaned with MilliQ water to remove dust particles. The cuvettes were then closed with a lid to avoid dust intrusion. The polydispersity was determined by dynamic light scattering. The instrument was equipped with one detector placed at an angle of 173°. Each result was the average of 30 measurements and each individual experiment was conducted in triplicate. Since the three size distributions obtained did not differ from each other, the average was taken to characterise the polydispersity.

6.3.3 Self-interaction chromatography (SIC)

Purified and crude BHA were independently immobilised by EDC and NHS (Moffett et al., 1993; Sehgal & Vijay, 1994) to attach the surface-accessible carboxylic groups covalently to the primary amino-groups of Tosoh Biosep AF-Amino-650-M chromatographic particles (Tosoh Bioscience, Montgomeryville, PA, USA). A total of 0.5 mL settled particles were extensively washed with DI water and resuspended in 10 mL of 8 mg/mL BHA solution buffered with 50 mM HEPES at pH 8.5 to which 6 mg NHS and 150 mg EDC were added. The suspension was allowed to react overnight at room temperature under gentle agitation on a rotary mixer. The particles were collected by passing the suspension over a glass frit covered with a 0.22 µm PTFE membrane (Millipore Corp., Bedford, MA, USA) and subsequently washed with 200 mL of 50 mM HEPES buffer at pH 8.5. Immobilised protein concentrations were determined directly by measuring the protein concentration on the particles using a micro BCA assay (Pierce, Rockford, IL, USA) following the instructions of the manufacturer. The columns (3 x 50 mm borosilicate glass microcolumns, Bio-Chem Valve/Omnifit, Boonton, NJ, USA) were packed at a flow rate of 5 mL/min against gravity using an automated Pharmacia FPLC system (Amersham Pharmacia Biosciences, Uppsala, Sweden). The flow rate was subsequently kept below 0.5 mL/min to prevent the bed from settling further. The integrity of the column was verified by injecting 25 µL of 1% acetone, which typically led to a symmetric peak.

For all experiments, a buffer containing 10 mM MES, 10 mM HEPES and 10 mM boric acid (MHB-buffer) was used. All reagents were dissolved in ultra-pure water of 18.2 MΩcm resistivity (Millipore Corp., Bedford, MA, USA), filtered via 0.22 µm vacu-cup Supor membranes (Pall, East Hills, NY, USA) and purged with helium for about 1 minute before

use. Where indicated, 0.2 % (w/v) of the antimicrobial agent, Proxel LV, was added. Most of the SIC measurements were performed on an ÄKTA Purifier HPLC unit equipped with an A-900 autosampler (Amersham Pharmacia Biosciences, Uppsala, Sweden) at room temperature. SIC measurements in the presence of Proxel and temperature-controlled experiments were performed on a Waters Alliance system with temperature-controlled sample and column chamber connected to a 2488 multi-channel UV/vis and a 2414 RI detector (Waters Corp. Milford, MA, USA).

The injected samples were prepared exclusively from the crude BHA material. BHA solutions of various salt concentrations were pH adjusted, subsequently filtered via syringe-driven 0.22 µm membranes (Pall, East Hills, NY, USA) and then loaded into the autosampler. The system was programmed to randomly inject the protein solutions in triplicate after 10 column volumes of equilibration at the desired salt concentration. Injection concentrations of 1 to 2 mg/mL BHA and volumes of 20 µL were found to lead to reproducible results. After the protein was eluted, the salt concentration was increased to 1 M for 4 column volumes and decreased to 0 M salt concentration for 7 column volumes before it was adjusted to the next salt concentration of interest. Since the peaks were usually symmetric, the peak maximum was used to determine the retention time. In spite of the aggregates present in the mobile phase (Figure 6.1), no separate peak was resolvable from the monomer peak. When not in operation, the columns were stored at 4°C in MHB-buffer at pH 9 without any preservatives like Proxel LV or sodium azide.

The chromatographic retention times are given in terms of the retention factor

$$k' = \frac{V_r - V_0}{V_0} \tag{13}$$

where V_r is the retention volume of the solute passing through the column under given solution conditions; V_0 refers to the retention volume under solution conditions at which the solute does not interact with the particle surface. V_0 was determined by passing the solute through columns filled with blank chromatographic particles, i.e., particles which have undergone the same chemical treatment as described above but in the absence of protein. B_{22} was calculated according to

$$B_{22} = B_{22}^{HS} - \frac{k'}{\rho_s \phi} = \frac{16}{3}\pi r^3 - \frac{k'}{\rho_s \phi} \quad (14)$$

B_{22}^{HS} is the excluded volume contribution, and ρ_s is the amount of immobilised protein per accessible surface area of the particle (Tessier et al., 2002a). The phase ratio ϕ is the total accessible surface area of the pore space per unit volume of the mobile phase, which is a characteristic of the chromatographic resin used (DePhillips & Lenhoff, 2000). The product $\rho_s\phi$ is a measure of the immobilised protein concentration available to interact with the mobile protein molecules. The equivalent radius r of the protein was determined from correlations for globular soluble proteins based on their molecular weight (Neal et al., 1998; Neal & Lenhoff, 1995). BHA molecules were assumed to be spherical with a monomer mass of 55 kDa, or an equivalent radius of 2.47 nm, and the corresponding excluded volume contribution B_{22}^{HS} was 252.5 nm^3 per BHA molecule. The amount of immobilised BHA was between 14 and 21 mg/mL settled particles. For BHA molecules in the Tosoh Biosep AF-Amino-650-M particles ϕ was estimated to be 15.9 m^2/mL mobile phase (DePhillips & Lenhoff, 2000).

6.4 Results and discussion

6.4.1 Polydispersity of BHA

This study aimed to evaluate SIC to determine the B_{22} values in polydisperse systems like BHA. In Figure 6.1, the polydispersity of BHA in MHB-buffer without additional salts but with and without the antimicrobial agent Proxel LV (0.2% w/v) is depicted. Apart from the monomer, which has a diameter of approximately 6 nm and accounts for the majority of protein molecules, two more populations were present, a smaller one at a diameter of 40 nm, and a bigger one at around 300 nm. Estimated from the intensities at the mode diameters of the individual populations, approximately 60% of the scattering intensity, which is the basis for B_{22} measurements by SLS, originated from the monomers. The addition of Proxel shifted

the population distribution towards the aggregates; the number of monomers decreased and the second population was enhanced in both number and size (now 50 nm). The third population stayed almost constant in number but had a reduced size (200 nm). In this case, approximately 50% of the scattering intensity was generated from the monomers. Naturally, B_{22} values determined by SLS become significantly less accurate for systems in which only between 50 and 60% of the scattering intensity originates from the target of interest, namely the monomers. According to Rayleigh's law, the scattering intensity varies strongly with the radius of the scatterer, i.e., $I \sim r^6$ (Wyatt, 1993). Consequently, the number of BHA aggregates was quite small and should therefore not have disturbed SIC measurements appreciably. Indeed, the peaks obtained from SIC runs were of high symmetry and no separate peaks other than the monomer were resolvable, which confirmed the assumption stated above (data not shown).

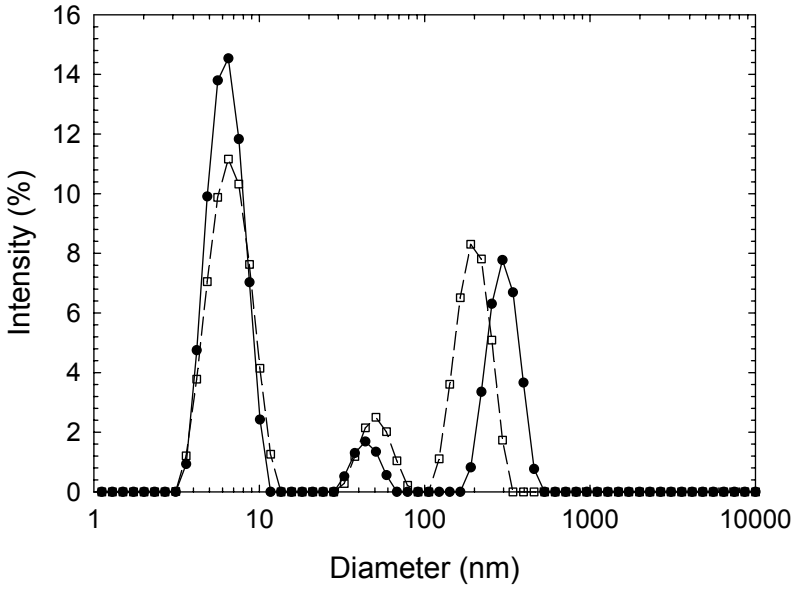

Figure 6.1 Polydispersity of BHA at pH 9 buffered in MHB; (●, solid line): no salt added; (□, dashed line): no salt added, 0.2% Proxel LV (w/v).

6.4.2 Influence of anions on the protein-protein-interactions

Since a large amount of solubility data for BHA as a function of anions and cations from the Hofmeister series are available and since one core goal of this study was to estimate how strong a link between B_{22} and the solubility can be established, the SIC experiments were conducted under the same conditions as the crystallisation experiments (Chapter 2), except for the temperature and Proxel (room temperature and no Proxel in SIC, but 40°C with Proxel (0.2% w/v) in solubility studies).

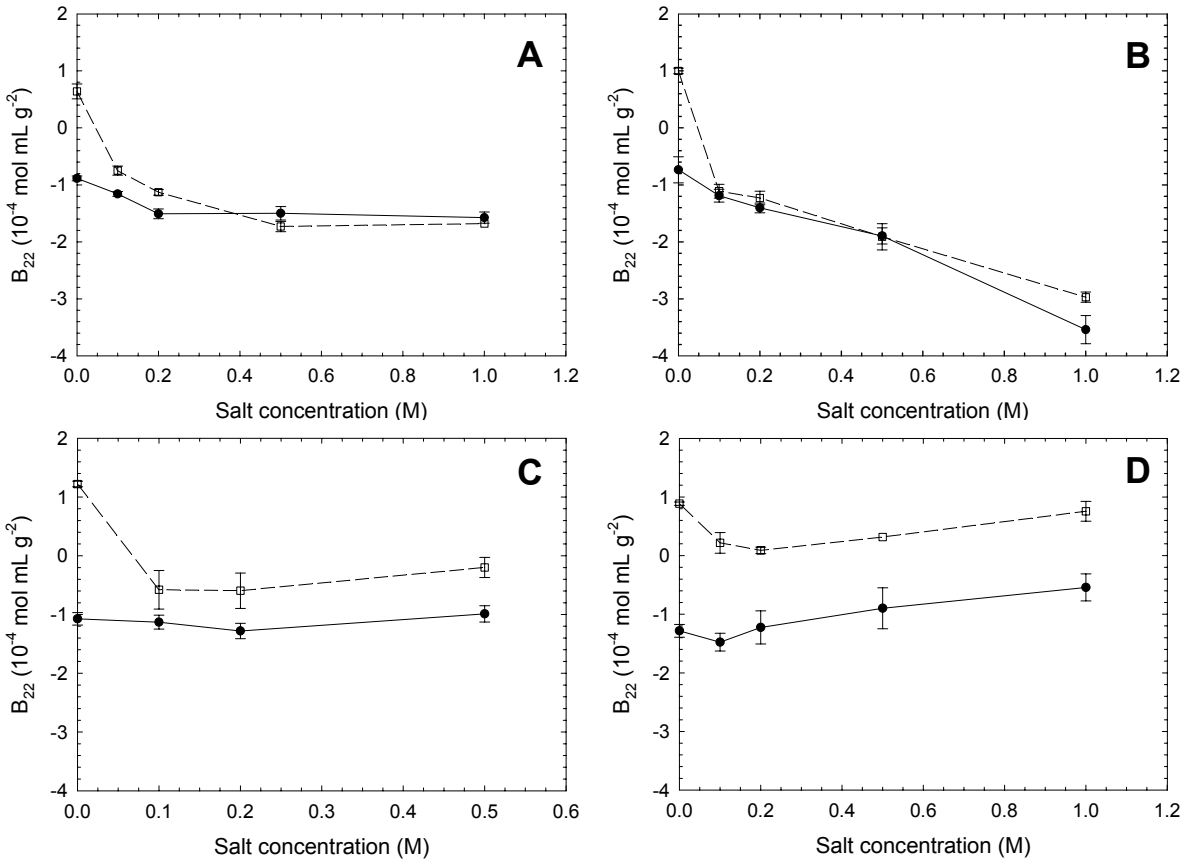

Figure 6.2 **B_{22} of purified (●, solid line) and crude (□, dashed line) BHA in MHB-buffer at pH 9 as a function of salt concentration. A: sodium chloride; B: sodium acetate; C: sodium nitrate; D: sodium thiocyantate.**

In Figure 6.2 A the influence of sodium chloride on the B_{22} of BHA is shown. The addition of sodium chloride led to a decrease in B_{22} up to 0.2 M for the purified and 0.5 M for the crude enzyme. Increasing the salt concentration up to 1 M had no significant additional effect on B_{22}. The differences in B_{22} obtained from the purified and crude immobilised material were pronounced at low salt concentrations but diminished with increasing salt concentrations. For the conditions tested, crystals were obtained in solubility experiments. According to the concept of the crystallisation slot, the B_{22} values of protein solutions leading to crystals should be located between -0.5 and -8×10^{-4} mol mL g^{-2}. The guidelines of this concept were fulfilled for almost all conditions, except for the solution free of sodium chloride using the crude material immobilised. All of the remaining B_{22} values were situated near the upper limit of the crystallisation slot.

The addition of sodium acetate provoked a larger decrease in B_{22}, meaning that the protein-protein interactions became more attractive (Figure 6.2 B). The decrease in B_{22} continued throughout the concentration range examined. Differences between the crude and the purified immobilised material were obvious at zero sodium acetate concentrations but were within experimental error at concentrations above 0.1 M. Only the B_{22} value for the solution free of sodium acetate measured on crude immobilised BHA fell outside the crystallisation slot.

Sodium nitrate concentrations higher than 0.5 M interfered strongly with the UV signal so that the protein-protein-interactions could be characterised only at lower salt concentrations (Figure 6.2 C). B_{22} decreased up to 0.2 M but then increased when the salt concentration was raised to 0.5 M. For the crude immobilised material, the minimum in B_{22} was more obvious than for the purified enzyme. For the latter the changes in B_{22} were within the range of the standard deviations so that B_{22} was essentially constant with increasing salt concentration. B_{22} values determined for the crude immobilised BHA were very close to or just outside the boundaries of the crystallisation slot.

The behaviour of B_{22} as a function of sodium thiocyanate concentration is shown in Figure 6.2 D. In both cases, i.e., crude and purified BHA immobilised, a minimum in B_{22} was encountered at sodium thiocyanate concentrations between 0.1 and 0.2 M. None of the B_{22} values determined for the crude BHA immobilised fell inside the crystallisation slot. Where the purified BHA was immobilised, the B_{22} values obtained above 0.5 M sodium thiocyanate concentration closely approached the upper limits of the crystallisation slot.

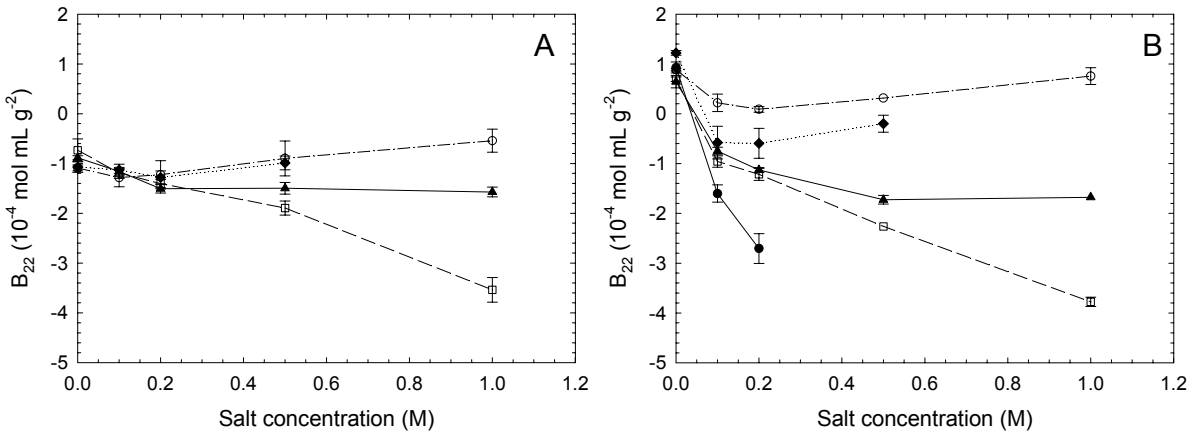

Figure 6.3 **A: B_{22} of the purified BHA in MHB-buffer at pH 9 as a function of the concentration of selected anions from the Hofmeister series. Sodium thiocyanate (○, dash-dotted line), sodium nitrate (◆, dotted line), sodium chloride (▲, solid line), sodium acetate (□, dashed line).**

B: B_{22} of the crude BHA in MHB-buffer at pH 9 as a function of the concentration of selected anions from the Hofmeister series. Sodium thiocyanate (○, dash-dotted line), sodium nitrate (◆, dotted line), sodium chloride (▲, solid line), sodium acetate (□, dashed line), sodium sulphate (●, solid line).

Figure 6.3 A and B summarise the B_{22} values measured for anions with the purified and the crude BHA immobilised, respectively. Where the purified BHA was immobilised, the effect of the different anions on B_{22} at concentrations of 0.2 M or lower was almost the same: they all decreased the B_{22} value slightly. Only above this concentration did each anion develop its own characteristic influence on the B_{22}. Acetate induced the strongest attractive protein-protein-interactions, followed by chloride. B_{22} values measured for nitrate and thiocyanate were very close to each other and led to less attractive interactions. The order in which the different anions affected the B_{22} follows the Hofmeister series (Hofmeister, 1888), which is in good agreement with the results of solubility experiments (Chapter 2). Where the crude BHA was immobilised, the same order was found. Here, the influence of sodium sulphate up to 0.2 M was included as well. Unfortunately, higher concentrations of sodium sulphate could not be measured because very broad peaks were obtained, probably caused by very strong interactions between the protein molecules in the stationary and the immobilised phase. As a consequence, measurements on the purified material using sulphate were not conducted.

Sulphates are known to be good precipitants, a fact which was reflected by low B_{22} values even at salt low concentrations, also in accordance with the Hofmeister series.

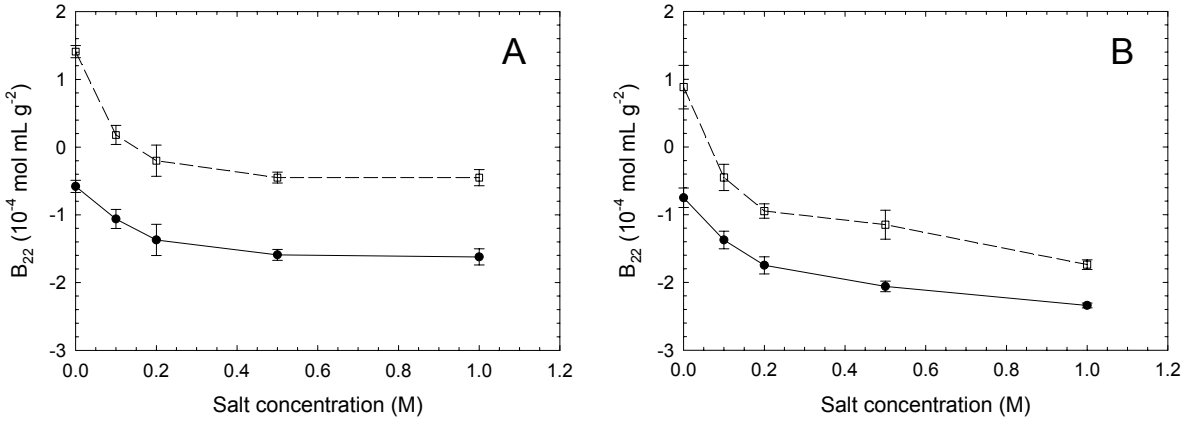

Figure 6.4 B_{22} of purified (●, solid line) and crude (□, dotted line) BHA buffered in MHB at pH 9 as a function of salt concentration. A: lithium chloride; B: potassium chloride.

6.4.3 Influence of cations on the protein-protein-interactions

The addition of lithium chloride caused a decrease in B_{22} up to ca. 0.5 M, but above this concentration B_{22} remained almost constant (Figure 6.4 A). The trends detected for the purified and the crude immobilised BHA were very similar, but the values for the purified immobilised BHA were lower by about $2x10^{-4}$ mol mL g^{-2} throughout the whole salt range examined. The B_{22} values obtained from the crude immobilised BHA were all above the crystallisation slot and those from the purified immobilised BHA were inside the slot.

The trends observed for potassium chloride (Figure 6.4 B) follow those for lithium chloride in that the values for the purified immobilised material were consistently about $2x10^{-4}$ mol mL g^{-2} lower that for the crude immobilised BHA. However, B_{22} continued to decrease with increasing potassium chloride concentration across the whole salt concentration range examined. All B_{22} values measured for the purified immobilised material with potassium chloride were located within the crystallisation slot, but those determined for the crude immobilised BHA entered the slot at potassium chloride concentrations above 0.2 M.

The consistent difference between results on the two columns for lithium and potassium chloride may suggest that systematic differences between the two columns, e.g., in immobilisation density or column packing, were responsible. However, the less systematic differences seen under other conditions for which results were presented earlier make this unlikely. In particular, the trends observed for B_{22} as a function of sodium chloride, described above (Figure 6.2 A), vary less consistently across the salt concentration range.

The influence of cations on B_{22} measured with purified immobilised BHA is summarised in Figure 6.5 A. Lithium induced the weakest interactions, followed by sodium and potassium. The series is in agreement with previous reports on the Hofmeister series for cations (Riès-Kautt & Ducruix, 1997). The different cations could not induce their own characteristic influence on the B_{22}. The influence of the anion, here chloride, was apparently more significant for the cations examined in this study. In all cases, B_{22} decreased up to at least 0.5 M salt concentration, and in the case of sodium and potassium chloride it continued to decrease with salt concentration at higher concentrations. All B_{22} values fell into the crystallisation slot.

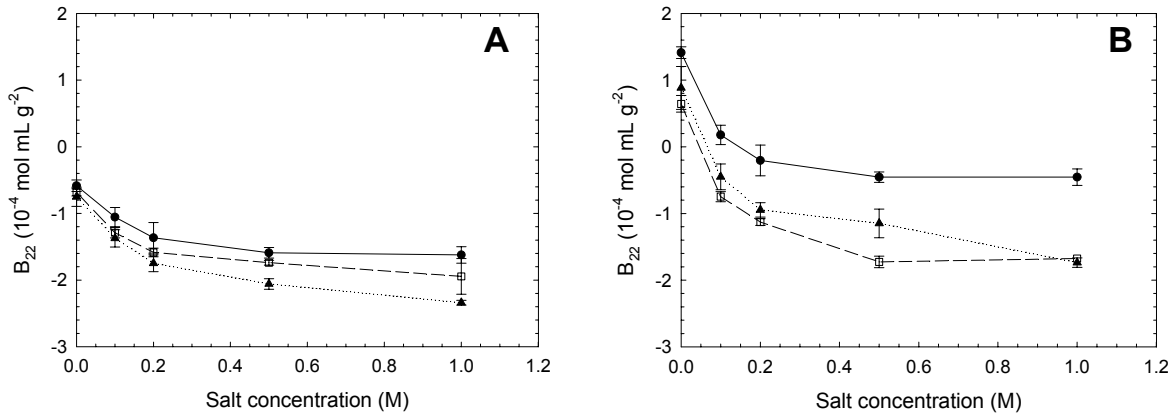

Figure 6.5 B_{22} of the BHA in MHB-buffer at pH 9 as a function of the concentration of selected cations from the Hofmeister series. Lithium chloride (●, solid line), sodium chloride (□, dashed line) and potassium chloride (▲, dotted line). A: purified BHA; B: crude BHA.

6.4.4 The role of impurities on the protein-interactions

Impurities can affect the measurements of the B_{22} in very different ways. They may e.g., block the pores of the chromatographic particles and may influence the determination of the protein surface density. Assuming that the impurities only influenced the protein-protein-interactions, the fact that the purified as well as the crude BHA were immobilised on the chromatographic resins may allow for an estimation of the influence of impurities on the protein-protein interactions.

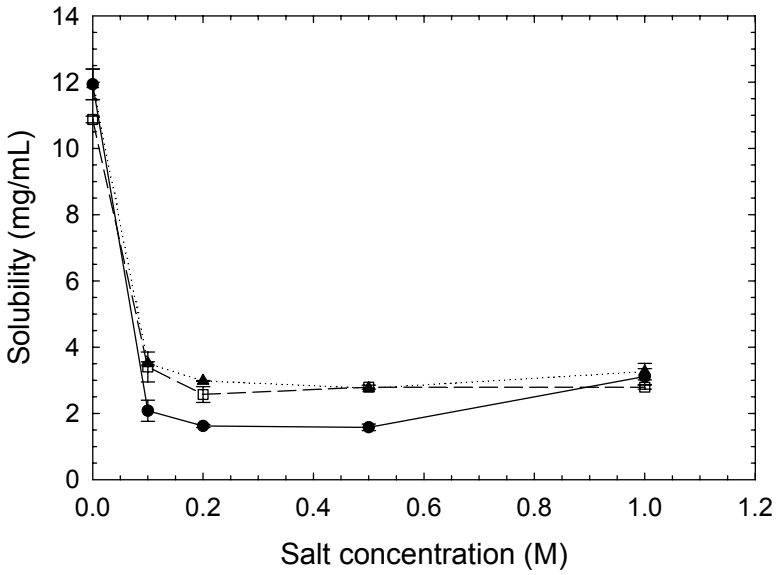

Figure 6.6 BHA solubility in MHB-buffer at pH 9 as a function of the concentration of selected cations from the Hofmeister series. Lithium chloride (●, solid line), sodium chloride (□, dashed line) and potassium chloride (▲, dotted line).

The use of columns prepared with the purified as well as the crude BHA allowed estimation of the influence of impurities on the protein-protein interactions. In all cases, B_{22} values obtained on purified immobilised BHA were lower, indicating that the interactions with the impurities were less attractive than BHA-BHA interactions. The differences in the B_{22} patterns with increasing salt concentration between the purified and the crude immobilised material can be subdivided into two groups: (i) the differences in B_{22} were higher at low salt concentrations but decreased with salt concentration and (ii) the differences were almost

constant throughout the salt concentration range examined. Salts showing pattern (i) behavior were sodium chloride and sodium acetate, whereas sodium nitrate, sodium thiocyanate, lithium chloride and potassium chloride gave rise to pattern (ii).

6.4.5 pH-dependence of B_{22}, using purified immobilised BHA

In Figure 6.7 A the pH dependence of B_{22} of the purified immobilised BHA at 0.1 M sodium chloride is presented for pH values between 8 and 10. Between pH 8 and 9, B_{22} increased from approximately -4 to -1.5x10^{-4} mol mL g^{-2} whereas it increased only slightly further from -1.5 to -1x10^{-4} mol mL g^{-2} when the pH was further increased to 10. This trend may be unexpected as the solubility showed a strong pH dependence throughout the same pH range in the absence of salts, with changes being even more pronounced between 9 and 10 than between 8 and 9 (Chapter 2). To further elaborate on the link between B_{22} and S, the solubility of BHA under identical conditions was determined, i.e., 0.1 M sodium chloride and a pH range between 8 and 10. The corresponding B_{22}-S pairs are given in Figure 6.7 B. This graph clearly demonstrates the agreement with typically observed trends between B_{22} and S (Guo et al., 1999) and underlines the above mentioned high sensitivity of the solubility to minor changes in B_{22} at the upper end of the crystallisation slot.

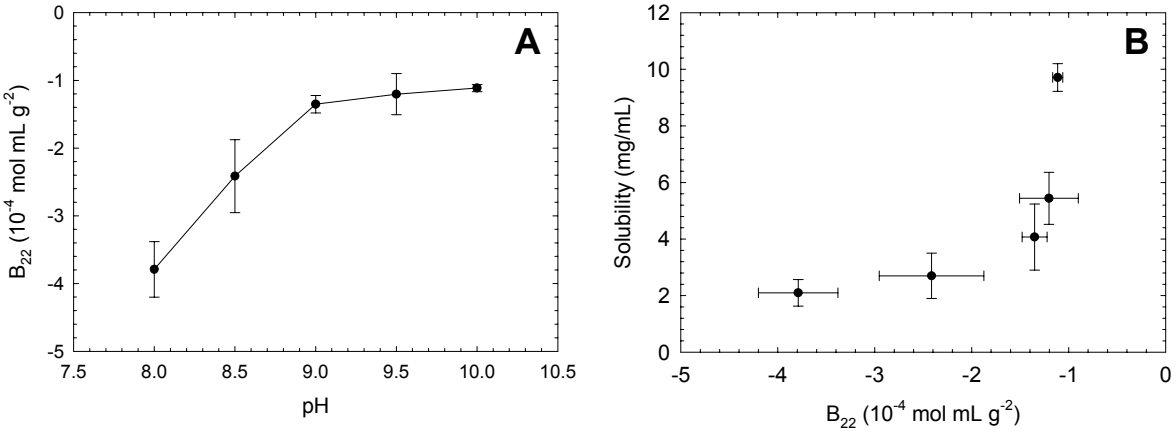

Figure 6.7 **A: pH-dependence of the B_{22} of purified BHA buffered in MHB at 0.1 M sodium chloride. B: Solubility over B_{22} of BHA; data recorded at pH-values between 8 and 10 at 0.1 M sodium chloride concentration.**

6.4.6 Link between solubility S and B_{22}

The Haas-Drenth-Wilson theory (HDW) suggests that the B_{22}-S pairs (i.e. B_{22} and S determined under identical conditions) of a given protein follow a characteristic trend that is captured by an idealised model of molecular interactions that they developed (Haas et al., 1999). A plot summarising all available B_{22}-S pairs, with B_{22} values measured on the purified BHA immobilised, is shown in Figure 6.8. The B_{22}-S pairs follow the typically observed trend that solutions of lower B_{22} exhibit lower solubilities and vice versa. Moreover, the sensitivity of the solubility to small changes in B_{22} in the upper part of the crystallisation slot is correctly represented. Although the scatter in the data is quite significant, particularly in the range of -1 to -1.4×10^{-4} mol mL g^{-2}, the development of the HDW model was based on B_{22}-S data available for lysozyme that included experimental scatter comparable to those found in the current study (Haas et al., 1999).

The data presented in Figure 6.8 were fitted by means of the HDW-model according to equation (15).

$$B_{22} = \frac{4}{M\rho}\left[1 - A\left\{\left(\frac{S}{m}\right)^{-\frac{2}{z}} - 1\right\}\right] \tag{15}$$

Here, M is the molecular weight of the protein (55 kDa); ρ is the density, which is typically set to 1.36×10^{3} kg m^{-3} for proteins and

$$m = \frac{M}{18\rho} \tag{16}$$

The HDW model contains two adjustable parameters. The first is z, the coordination number, which represents the number of nearest neighbouring protein molecules inside the crystal lattice and usually depends on the crystal structure and the packing fraction. Alternatively, z can be interpreted as the number of macro-bonds in the crystal lattice. The other parameter, A, depends on the anisotropy p (p = 1 for isotropic interactions and p < 1 for anisotropic interactions) and on the range of interactions between protein molecules. A is characteristic of

each individual protein and can take any value, but it is small for strong anisotropy and short-range interactions, which are typical of real protein interactions.

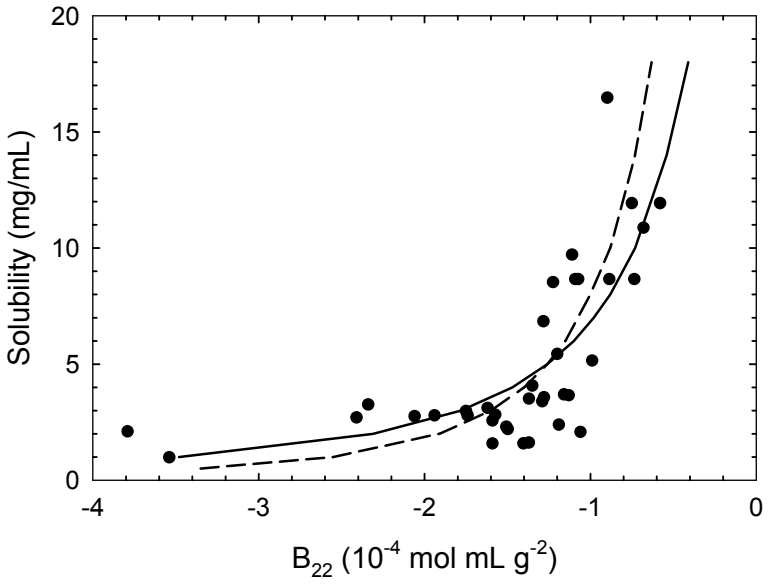

Figure 6.8 Solubility versus second osmotic virial coefficient. Summary of all available S-B_{22} pairs (B_{22}-values obtained on the purified immobilised BHA), regardless of the salt type and concentration (●). HDW-model fitted with the parameters z = 6 and A = 0.04 (dashed line) and z = 4 and A = 0.0043 (solid line)

Table 6.1 Influence of the choice of different coordination numbers on the anisotropies p of BHA

Coordination number z	6	4
A	0.04	0.0043
range of interaction (Å)	***p***	***p***
1	0.787	0.085
2	0.387	0.042
3	0.254	0.027

Fits are shown in Figure 6.8 for z = 6, A = 0.04, and for z = 4, A = 0.0043. Neither fit clearly represents the experimental data better within the scatter. The calculated anisotropies corresponding to the two different coordination numbers at ranges of interactions of 1, 2 and 3 Å differ appreciably (Table 6.1). The corresponding variation in the adjustable parameter A is about one order of magnitude. Within the range of interaction of between 1 and 3 Å, p values of between 0.1 and 0.03 respectively were calculated for lysozyme for z = 4 (Haas et al., 1999), and at ranges of 2 and 3 Å, values of 0.89 and 0.58 respectively were found for equine serum albumin, and of 0.75 and 0.49 respectively for ovalbumin for z = 6 (Demoruelle et al., 2002). The reason that p-values for equine serum albumin and ovalbumin are not given at a distance of 1 Å is that the calculations would yield p-values higher than 1, which is impossible. However, these findings are in qualitative agreement with the current study in that the anisotropy of the interactions is very sensitive to the choice of the coordination number. Equine serum albumin has a molecular weight (MW) of 65.5 kDa and a hydrodynamic radius (r_h) of 3.8 nm, whereas the MW of ovalbumin is 45 kDa and the r_h 2.8 nm. Judging from these units the two proteins are not that different from BHA (MW = 55 kDa and r_h = 3 nm). However, more isotropic interactions, i.e. higher p-values, were suggested for equine serum albumin and ovalbumin compared to BHA if z was chosen to be 6. The comparison between lysozyme and BHA at z = 4 revealed very similar p-values which may be interpreted in a way that the interactions between two molecules of these two proteins have similar characteristics, a fact which will be further discussed later.

The B_{22}-S-pairs presented in the current study did not lead to a clear conclusion which coordination number was the correct one. It is believed that a better data quality is the key for obtaining more sound values for the anisotropy of the interactions.

6.4.7 Influence of antimicrobial agents on protein-protein interactions

Influence of Proxel on B_{22}

The experiments to determine BHA solubility lasted 96 hours for each condition, which is short in comparison to methods employed for other proteins (Ataka & Tanaka, 1986). However, microbial contamination cannot be excluded even within this relatively short time. Since the influence of microbial contaminations on crystal growth and solubility is unknown, an antimicrobial agent, namely Proxel LV, was always added to exclude any uncertainty. Since it was present in relatively small amounts (0.2 % w/v), the influence of Proxel LV on the solubility was never explicitly examined and assumed to be negligible. However, in the B_{22} measurements on BHA it was found that the presence of this agent caused an extremely strong absorbance at wavelengths typically employed to record the protein peaks, i.e., 300 nm and below. As a consequence, all the B_{22} values presented in the current study were determined in the absence of Proxel. In contrast to the solubility determination, the measurement of a complete set of B_{22} values by SIC took less than 12 hours, so that the risk of influential microbial contaminations could be excluded. However, in relating B_{22} to solubility it is imperative that the B_{22}-S pairs are obtained under identical solution conditions.

To address this conundrum SIC was performed in the presence of Proxel using the Waters HPLC system connected to an RI detector. The determination of B_{22} values was very challenging, however, mainly because of unstable RI signals, so the data obtained are subject to large uncertainties. The results presented in Figure 6.9 can be summarised as: (i) the RI detector seemed to yield lower B_{22} values than the UV detector and (ii) the B_{22} values obtained in the presence of Proxel were lower than without. It may be concluded from these results that Proxel lowered the B_{22} value, i.e., it induced more attractive protein-protein interactions.

To minimise the uncertainty inherent to the B_{22} values based on the RI detector, solubility measurements were also conducted with and without Proxel. The results are presented in Figure 6.10, which also shows the influence of 10 mM sodium azide, as discussed later. The addition of 0.2 % (w/v) Proxel did not have a noticeable influence on the solubility of BHA. It can thus be concluded that the overall influence of Proxel on B_{22} is probably small, consistent with the implicit assumption of a negligible influence of Proxel on the solubility. Therefore,

although S and B_{22} were recorded with and without Proxel, respectively, the correlation of these two units by means of the HDW model was justifiable.

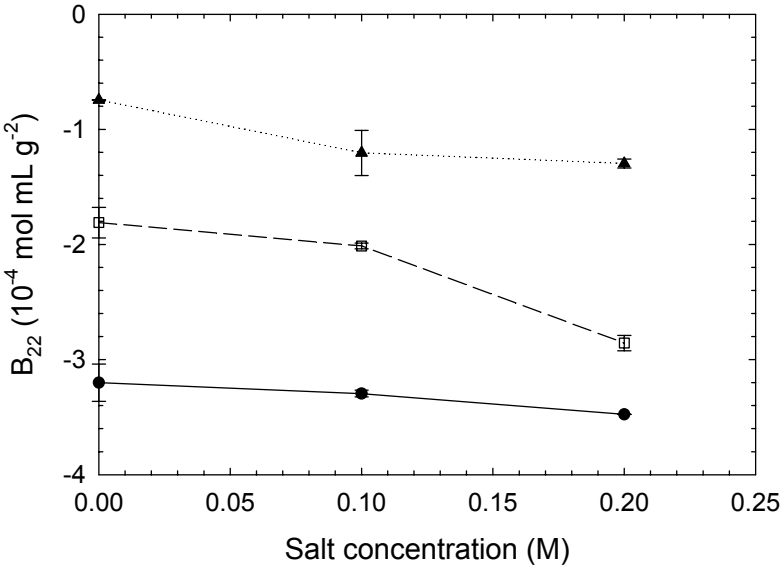

Figure 6.9 Comparison of the B_{22}-value as a function of sodium chloride concentration in the presence of 0.2% (w/v) Proxel (●, solid line, RI signal), no Proxel (□, dashed line, RI-signal) and no Proxel (▲, dotted line, UV-signal).

The effect of sodium azide on the protein-protein interactions

Since severe problems in the determination of the B_{22} in the presence of Proxel were encountered, the use of alternative antimicrobial agents was considered. Probably the most commonly employed agent is sodium azide, which was shown not to absorb at a wavelength of 280 nm in the concentration ranges typically used to take advantage of its germicidal effect (i.e. 5 to 30 mM, data not shown). The influence of 10 mM sodium azide on the B_{22} of BHA as a function of potassium chloride was examined (Figure 6.11), in testing the purified immobilised BHA. At zero salt concentration, almost no differences in B_{22} were found. Up to 0.2 M potassium chloride concentration, B_{22} in the azide-free solution decreased more rapidly than with 10 mM sodium azide. At higher potassium chloride concentrations, the two B_{22} values decreased to almost the same degree. In this concentration range, B_{22} of the azide-

containing solution was increased by approximately 10^{-4} mol mL g^{-2}. Such behaviour was not seen when the influence of impurities on B_{22} was examined. Interestingly, no difference in B_{22} was found when no potassium chloride was present. This finding would be in agreement with the solubility measurements and suggests that azide in the tested concentration range did not impose any influences on the protein interactions (Figure 6.10). However, once the increased potassium chloride concentration imposed more attractive protein interactions, sodium azide seemed to have a moderating effect, leading to less negative B_{22} values. This leads to the conclusion that interactions between potassium chloride and sodium azide exhibited a decisive influence on the protein interactions. Azides are known to be chaotropes (Stirpe et al., 2002) which would be in line with the finding that their presence leads to less attractive protein-interactions. Thus, a solubilising effect would be expected which was, however, not confirmed by solubility measurements. Reasons for these discrepancies are not known.

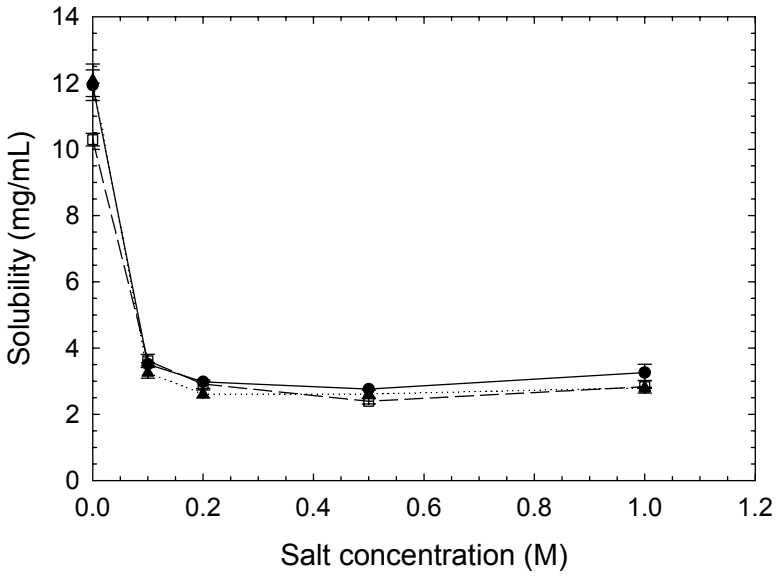

Figure 6.10 Influence of antimicrobial agents on the solubility of BHA at pH 9, 40°C as a function of potassium chloride concentration. 0.2 % (w/v) Proxel, no azide (●, solid line); no Proxel, no azide (□, dashed line); no Proxel, 10 mM sodium azide (▲, dotted line).

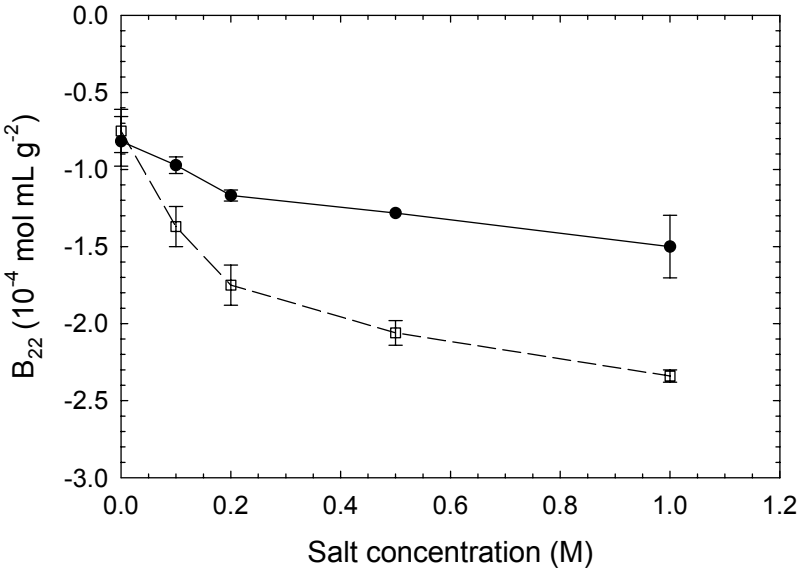

Figure 6.11 Influence of 10 mM sodium azide on the B_{22} of purified immobilised BHA at pH 9, as a function of potassium chloride concentration. 10 mM sodium azide (●, solid line); no sodium azide (□, dashed line).

6.4.8 The role of the temperature on B_{22}

As discussed above, meaningful links between B_{22} and S can only be based on B_{22}-S pairs measured under identical conditions. Solubility trials were conducted at a temperature of 40°C whereas SIC measurements were carried out at room temperature, i.e., ca. 25°C. It may be argued that the temperature has an influence on solubility and thus also on B_{22}. Some proteins like carbomonoxy-hemoglobin C (Vekilov et al., 2002) possess a strong retrograde solubility dependence on temperature whereas the opposite is true for lysozyme (Pusey & Munson, 1991). It has previously been shown that the temperature dependence of the BHA solubility at pH 9 between 25 and 40°C is negligible (Chapter 2). B_{22} experiments on the BHA system conducted at different temperatures (column and sample temperature controlled) have also demonstrated the temperature-independence of B_{22} (data not shown). It is therefore permissible to correlate B_{22} to S, albeit they were not obtained at the same temperature.

In spite of the discrepancies encountered between solubility and B_{22} measurements as discussed above, along with the sometimes large experimental scatter of the B_{22}-S pairs, it is

argued that the correlation found via the HDW model is still of acceptable quality. Apart from the fact that the B_{22}-S pairs were not obtained under identical conditions (in terms of temperature and antimicrobial agent), the experiments were conducted up to one year apart in different laboratories. It is believed that the data quality could be improved significantly by running synchronised experiments in place and time to maximise similarities in the conditions. However, the correlation found between B_{22} and S may still be beneficially employed to evaluate the influences that additives of interest may have on solubility. In many cases, the prediction of the correct trend may be more important than obtaining exact numerical values for the protein solubility.

6.4.9 Comparison of the BHA self-interactions with other well-characterised proteins

The large amount of solubility and B_{22} data available enables an extensive comparison of BHA to other well-characterised proteins. Three main phenotypes of protein-protein interactions as a function of salt concentration have previously been described (Tessier et al., 2002b).

Group 1: This group of proteins is characterised by a reduced intrinsic affinity to self-associate, even at high salt concentrations, meaning that they display repulsive interactions throughout a large range of salt concentrations, reflected by positive B_{22} values. Only at relatively high salt concentrations, typically above 1.5 M, is a sharp drop in B_{22} observed. In many cases B_{22} passes through the crystallisation slot within a very narrow window of salt concentrations. Examples of this phenotype are bovine serum albumin (BSA), myoglobin (Tessier et al., 2002a) and calcium- and integrin-binding protein (CIB) (Berger et al., 2005).

Group 2: Proteins belonging to this group display B_{22} values that decrease with increasing salt concentration as well as with approach of the solution pH to the protein's pI. The behaviour is similar to that described by the DLVO theory and emphasises the role of electrostatic interactions dominated by the net charge of the protein. Example of proteins belonging to this group are lysozyme (Tessier et al., 2002a) and subtilisin (Pan & Glatz, 2003).

Group 3: Depending on the pH, proteins belonging to this group show attractive interactions at low ionic strength even if their net charge is different from zero. This behaviour has been

ascribed to chymotripsinogen (Velev et al., 1998) and ribonuclease A (Tessier et al., 2003). Reasons for this behaviour could be complex dependencies on the orientation of protein interactions (Neal et al., 1998).

B_{22} values (Tessier et al., 2002a) and solubilities (Broide et al., 1996) passing through a minimum with increasing magnesium bromide concentrations have been reported for lysozyme. The absence of salting-in but presence of a pronounced salting-out effect at the addition of sodium chloride has been demonstrated by solubility (Retailleau et al., 1997) and B_{22} measurements (Tessier et al., 2002a), for BSA at pH 7 (Tessier et al., 2002b) and subtilisin at pH 5.5. Salting-out of the latter was also seen with sodium thiocyanate (Pan & Glatz, 2003). A minimum in B_{22} at 0.15 M sodium acetate was measured for apoferritin (Petsev & Vekilov, 2000) and a solubility minimum was found for xylose isomerase (Vuolanto et al., 2003) at 0.15 M magnesium sulphate. When the pH of lysozyme solutions (pI ~ 11) was decreased from 10.5 to 3, B_{22} increased monotonically and became positive at very low pH values and salt concentrations (Velev et al., 1998).

When the B_{22} trends for BHA are compared with those for the three groups described above, it becomes obvious that BHA shares characteristic properties with subtilisin, apoferritin, xylose isomerase and lysozyme. BHA and lysozyme exhibit the same pH dependence and the absence of a salting-out effect. Moreover, a minimum in B_{22} with increasing concentrations of particular salts has been observed for both proteins. The minimum was found to be at similar salt concentrations (0.2 M for BHA and 0.3 M for lysozyme), although different conditions and salts were employed. The B_{22} minimum for apoferritin (Petsev & Vekilov, 2000) and the solubility minimum for xylose isomerase (Vuolanto et al., 2003) were in a similar salt concentration range (0.15 M in both cases). Interestingly, the anisotropy of interactions within a distance of 1 to 3 Å was similar for lysozyme and BHA when the coordination number was chosen as 4. Lysozyme and BHA are not the only exceptions to the broadly accepted idea of protein interaction behaviour that results in increasing solubility with increasing salt concentration (salting-in) up to a certain concentration followed by a decrease (salting-out) (Edsall, 1930; Green, 1932) that is ascribed to changes in the chemical potential of the protein imposed by the salts via hydrophobic effects. The current study emphasises the dominant role of electrostatic forces on protein interactions, which is in agreement with the findings of solubility and zeta potential studies conducted on BHA (Chapter 2) and BLA (Chapter 4).

6.5 Conclusions

Measurements of the second osmotic virial coefficient B_{22} are typically employed to identify conditions leading to the growth of high quality protein crystals starting from highly purified material suitable for X-ray diffraction studies. In this study the application of B_{22} was extended to rapidly obtain information on how different additives like salts influence the solubility of semipure preparations of the industrially relevant recombinant *Bacillus halmapalus* α-amylase (BHA). B_{22} was determined by self-interaction chromatography (SIC), which has been shown to be robust to the presence of soluble aggregates and impurities and offers the advantage of distinctly reduced time and protein demands. The anisotropies of the interactions within the framework of the Haas-Drenth-Wilson (HDW) model as well as the trends of B_{22} as a function of pH and salt concentration revealed similarities between BHA and lysozyme. B_{22} was linked to the solubility by the HDW model. The experimentally obtained B_{22}-S pairs qualitatively followed the characteristic trend suggested by the HDW model, although significant data scatter reduces the quantitative rigor of the correlation between B_{22} and S. The biggest discrepancies between S and B_{22} were found at high salt concentrations (0.5 to 1 M potassium chloride) with the antimicrobial agent sodium azide added in relatively small amounts (10 mM). Moreover, SIC and solubility measurements suggested different positions of lithium within the Hofmeister series for cations. In spite of uncertainties in the correlation between B_{22} and S, the approach pursued in the current study is believed to be promising since SIC was shown to be able to replace tedious solubility experiments and to rapidly provide insight into aggregation and crystallisation processes, which can rationalise the search for novel proteins and protein formulations with advantageous properties. SIC may also emerge as an indispensable tool in the design, control and effective operation of production and recovery processes for tailoring any given protein. The widespread applicability, its robustness towards less pure systems and the ease of operation and automation can make SIC into a valuable instrument to be beneficially employed in very different stages of protein discovery and production projects and holds promise as a powerful method for process control as well as in industrial and academic protein research.

7 Final conclusions and further perspectives

In this final chapter, the tools and methods employed in the experimental part of this thesis are evaluated for their applicability in protein production. Moreover, their potential for characterising enzymes regarding their phase behaviour and self-interactions in view of the findings of this thesis is demonstrated.

7.1 Evaluation of the tools developed and applied in this thesis

The overall aim of this thesis was to evaluate traditional and novel approaches for the characterisation and quantification of the influence of different solution properties such as changes in salt concentration, pH and temperature on α-amylase solubility, crystal growth and the overall phase diagram. The main tools employed were traditional batch crystallisations conducted in Eppendorf tubes and in microtitre plates, the zeta potential and self-interaction chromatography. All of the tools showed high robustness with regard to impurities and fulfilled the requirement of reduced protein consumption, although to a very different extent. For the determination of a solubility curve which e.g., consisted of five different salt concentrations typically required 700 mg of protein and around 1300 mg was needed for the generation of an entire phase diagram as presented in this thesis. The determination of the second osmotic virial coefficient at five different salt concentrations reduced the protein consumption to approximately 2 mg but it has to be considered that a certain amount of protein is required to be immobilised to the chromatographic resins. A surface coverage of approximately 20 mg/mL settled chromatographic particles was used which means that for the packing of the column as used in this study (V = 0.4 mL) a minimum of approximately 8 mg of protein is needed. In practice though, an amount of 40 mg of protein is probably realistic to take account possible protein loss during the immobilisation procedure. However, the protein has to be immobilised only once and the packed column can be used for the screening of a large variety of conditions. When comparing the B_{22} measured for the same condition when the column had been used for B_{22}-determinations of 40 other conditions in between, the relative standard deviations were within 10%.

The choice of methods available for the characterisation of protein interactions will depend on a number of factors, such as the number of enzymes or enzyme variants to be analysed, the amount of material available, the purity of the enzyme to be analysed, and the number of conditions to be tested. Early development phases are typically characterised by the need to determine solubility behaviours of many variants of limited purity and availability in low concentration for a few conditions only. In this development phase, self-interaction chromatography is regarded as the method of choice since it requires only very little protein which can be immobilised even in low concentrations and in the presence of impurities. Further miniaturisation of SIC to "lab-on-chip-scale" has been proven to be feasible which would reduce the protein demand even further, approximately 500-fold (García et al., 2003). Moreover, the microchip technology has the advantage that the loading of the microchip columns with chromatographic particles is automated. However, SIC is outdoing its strength when many different solution conditions are to be tested for a small number of enzyme variants. In contrast, when a few process conditions should be tested on many enzyme variants, as often the case in early development phases, SIC has the disadvantage that the immobilisation of the proteins to the resins and the determination of the resulting surface coverage are time-consuming steps which can to date not be automated and could be the limiting step in high throughput screening routines.

In later stages of the development processes only very few variants of beneficial properties are tested which are typically available in larger quantities and in higher concentrations. The number of conditions tested depends on stability limits of the protein and the constraints imposed by the production which could e.g., be the choice of chemicals and their concentration, processing temperatures and pH-values. Although the constraints narrow down the number of process relevant conditions, process optimisation may still require the screening of many conditions for the comprehensive description of the phase behaviour of a given candidate. In this step, the use of microtitre plates to estimate the enzyme's phase behaviour is recommendable. The characteristics of the phase diagram for a given promising enzyme variant should provide decision support for the production strategy. In case the enzyme variant is very soluble, a high-throughput production strategy completely in the liquid phase is appropriate. If the variant is generally of low solubility but can be crystallised under process relevant conditions, the crystallisation process should further be optimised. In this case, the knowledge of the solubility curves is now of importance which can be determined in Eppendorf tubes. In parallel, the impact of core process parameters such as supersaturation,

temperature, ionic strength and stirring rate on the crystallisation processes should be quantified by means of crystal concentration, size and distribution as well as lag phase and time to reach equilibrium. For this purpose, the electrical sensing zone method has been shown to be an appropriate measuring technique. Corresponding experiments could also be conducted in Eppendorf tubes. If necessary, further experiments based on factorial design could be conducted in microtitre plates in search for more applicative crystal morphologies.

Once optimal crystallisation conditions in mL-scale have been identified, the process has to be transferred to larger scale. Crystal concentration and size distributions are appropriate measures to evaluate the quality of the larger scale crystallisation processes. Already during the scale-up procedures, the aptitude of SIC as a tool for online B_{22} measurements has to be evaluated and a link established between B_{22} and solubility, e.g., by the Haas-Drenth-Wilson model. Simultaneous online measurements of the B_{22} and the protein concentration by SIC and UV-absorption, respectively, could be used to conduct the process within the desired supersaturation ranges which would help obtaining reproducible crystal sizes and distribution. Together with phase diagrams, the online control of the supersaturation would open up possibilities to employ new processing strategies such as seeding, constant supersaturation control or any other promising strategy which may lead to improvements in the final crystal habit.

7.2 Characterisation of enzymes by their phase behaviour and self-interactions

One main aim of this thesis was to evaluate tools to characterise two enzymes of limited purity by their solubility, crystallisation and phase behaviour. The tested methods could provide following information: Of all parameters tested, the pH had the largest impact on solubility. The pH of zero zeta potential and the pI determined by isoelectric focusing were in good agreement with each other and the zeta potential and the charge curve calculated on the amino acid sequence followed similar trends as a function of pH. Regardless of the type of salt added, the solubility decreased up to a concentration of between 0.1 and 0.2 M. At higher concentrations it depended on the salt whether the solubility either further decreased or raised again. The Hofmeister series for anions was followed in the correct order. With the exception of lithium the efficiency of cations to influence the solubility was reversed to what was

expected. A reversal of this series upon polarity change of the protein net charge could not conclusively be demonstrated. Temperature had only a minor impact on the protein interactions. Complete phase diagrams, which were generated for the first time, revealed that the enzyme studied crystallised under a large variety of conditions. The crystal morphology was independent of the pH and supersaturation as well as the presence of salt. Within the supersaturation range tested, very small crystals of a mean diameter of 5 μm were found. The crystal size distribution did not change during the crystallisation process but the crystal concentration increased until the equilibrium concentration was reached. Very high supersaturations were necessary to induce the formation of amorphous precipitation. The metastable zone was small as a function of pH and at pH 9 as a function of sodium chloride but increased with sodium chloride and sodium thiocyanate concentration at pH 7. The second osmotic virial coefficient followed the same trends as the solubility and a correlation between the two units could be established such that tedious crystallisation processes to determine the solubility curve can to a large extent be replaced by the time and protein saving measurements of the B_{22}. The experimental results suggested that protein-protein interactions and thus solubility of the tested enzymes were essentially governed by electrostatic interactions. The tools evaluated in this study were shown to provide detailed information on how the chemical solution conditions, i.e. mainly pH, salt type and concentration, should be changed to achieve the desired effects on the phase behaviour of a given enzyme.

7.3 Future perspectives

The tools and methods developed and applied during this work constitute a technology platform which represents a core contribution for the rapid, robust and material-saving characterisation of solution properties in terms of solubility, stability, crystal growth and overall phase diagrams. This technology platform is believed to create far-reaching impulses for substantially accelerating the entire launching process of novel biological products, ranging from early screening routines to production and final formulation in large scale. With the aid of the approaches presented in this thesis, alternative processing strategies may become applicable which to date fail due to limited knowledge of the protein phase behaviour and due to missing tools for online measurement and control of solubility and supersaturation.

8 Appendix I: Quantification of the kinetics of BHA batch crystallisation processes

8.1 Introduction

This study aims for the quantitative description of BHA batch crystallisation processes using classical theories of nucleation and growth. The concentration profiles and the crystal concentrations determined in CP01 and CP02 (Chapter 3) were incorporated into mass and population balances. Nucleation was considered to be homogeneous only. Screw dislocation growth controlled by integration was assumed.

8.1.1 Population and mass balances

The growth and nucleation rates are incorporated in population balances which can be expressed by moments as

$$\frac{d\mu_0}{dt} = \frac{J}{\rho_{water}} \tag{17}$$

$$\frac{d\mu_k}{dt} = k\,\mathrm{G}\,\mu_{\mathrm{k\text{-}1}} + \mathrm{J}\,\mathrm{d_c}^{k} \tag{18}$$

For k = 1, 2, 3 and 4

$$\mu_k = \int_0^{\infty} \mathrm{L}^{\mathrm{k}}\, n(L)\, dL \tag{19}$$

The mass balance gives

$$\frac{dC}{dt} = -\frac{dm_c}{dt} = -\alpha\,\rho_{\mathrm{c}}\,\frac{d\mu_3}{dt} \tag{20}$$

The equations are normalised as

$$\theta = \frac{G_0}{L_0} t;\ \mu_0^{\bullet} = \frac{\mu_0}{N_0};\ \mu_1^{\bullet} = \frac{\mu_1}{L_0};\ \mu_2^{\bullet} = \frac{\mu_2}{S_0};\ \mu_3^{\bullet} = \frac{\mu_3}{V_0};\ \mu_4^{\bullet} = \frac{\mu_4}{Q_0} \text{ and } \mathrm{S} = \frac{C}{C^*} \tag{21}$$

For the calculation of N_0, L_0, S_0, V_0, and Q_0, the total weight per kg of water that can be crystallized (C_i-C*) has been chosen as a reference:

$$N_0 = \frac{(C_i - C^*)10^{-3}}{\alpha\, \rho_c\, \mathrm{d}_{c0}{}^3} \tag{22}$$

Where d_{c0} is the diameter of the critical nucleus at the initial supersaturation ratio, and C_i is the initial concentration.

$$L_0 = N_0 d_{c0};\ S_0 = N_0 d_{c0}{}^2,\ V_0 = N_0 d_{c0}{}^3;\ Q_0 = N_0 d_{c0}{}^4 \tag{23}$$

$$\frac{d\mu_0^{\bullet}}{d\theta} = \frac{J}{\rho\, \mathrm{w}_{\text{water}}} \frac{L_0}{G_0 N_0} \tag{24}$$

$$\frac{d\mu_1^{\bullet}}{d\theta} = N_0 \left(\frac{\mathrm{G}}{\mathrm{G}_0} \mu_0^{\bullet} + \frac{\mathrm{d}_c}{\mathrm{L}_0} \frac{d\mu_0^{\bullet}}{d\theta} \right) \tag{25}$$

$$\frac{d\mu_2^{\bullet}}{d\theta} = N_0 \left(2 \frac{\mathrm{G}}{\mathrm{G}_0} \mu_1^{\bullet} + \frac{\mathrm{d}_c{}^2}{S_0} \frac{d\mu_0^{\bullet}}{d\theta} \right) \tag{26}$$

$$\frac{d\mu_3^{\bullet}}{d\theta} = N_0 \left(3 \frac{\mathrm{G}}{\mathrm{G}_0} \mu_2^{\bullet} + \frac{\mathrm{d}_c{}^3}{V_0} \frac{d\mu_0^{\bullet}}{d\theta} \right) \tag{27}$$

$$\frac{d\mu_4^{\bullet}}{d\theta} = N_0 \left(4 \frac{\mathrm{G}}{\mathrm{G}_0} \mu_3^{\bullet} + \frac{\mathrm{d}_c{}^4}{Q_0} \frac{d\mu_0^{\bullet}}{d\theta} \right) \tag{28}$$

Thus, the mass balance becomes

$$\frac{dS}{d\theta} = -(S_{init} - 1)\frac{d\mu_3^{\bullet}}{d\theta} \tag{29}$$

8.1.2 Nucleation and growth

In the calculations of nucleation and growth rates, the crystals are assumed to be of spherical shape. There are two main types of nucleation, namely homogeneous and heterogeneous nucleation. Here, nucleation is assumed to be at steady state and exclusively homogeneous, (Kashchiev & van Rosmalen, 2003). At steady-state, the source of the material to produce subcritical clusters is sufficient to replenish the material that is lost as a critical cluster becomes a stable precipitate and grows (Vekilov et al., 1996). The nucleation work for a spherical crystal W* and the number of molecules in the nucleus n* are given by

$$W^* = \frac{16\pi\upsilon_0^{\,2}\gamma^3}{3\left(\mathrm{k_B\,T}\right)^2 \ln^2 \mathrm{S}} = \frac{1}{2}\mathrm{n^*\,k_B}\,T\ln S \tag{30}$$

$$n^* = \frac{32\pi\upsilon_0^{\,2}\gamma^3}{3\left(\mathrm{k_B\,T}\right)^3 \ln^3 \mathrm{S}} \tag{31}$$

where k_B is the Boltzmann constant, T the absolute temperature in Kelvin, v_0 the volume occupied by one molecule in the cluster (m^3), γ the specific surface energy of the cluster/solution interface (J/m^2), and S is the supersaturation ratio. v_0 and γ characterise the nanoscopically small clusters and are approximated by the classical nucleation theory as corresponding to macroscopically large crystals as

$$v_0 = \frac{\pi}{6} d_0^3 \tag{32}$$

where d_0 is the diameter of the protein molecule. The critical diameter of the nucleus is given by

$$d_c = \frac{\beta}{3\alpha} \frac{2 \upsilon_0 \gamma}{k_B T \ln S} \tag{33}$$

The nucleation kinetics can be calculated as

$$J = z f^* C_0 \exp\left(-\frac{W^*}{kT}\right) \tag{34}$$

with the Zeldovich or compressibility factor z

$$z = \left(\frac{W^*}{3\pi k_B T (n^*)^2}\right)^{1/2} \tag{35}$$

and f* which is the monomer attachment frequency which is controlled by mass and/or heat transfer. Two mechanisms of mass transport exist in the liquid phase.

8.1.3 Diffusion

Diffusion of molecules towards the clusters can take place either in the volume of the liquid phase or along the surface of a substrate. The attachment controlled by volume diffusion can be described as

$$f^*_{\text{diffusion}} = \left(48\pi^2 \nu_0\right)^{1/3} D\, C'_{eq}\, S\, n^{*1/3} \tag{36}$$

D is the diffusion coefficient of a protein molecule in the solution and is assumed equal to the diffusion coefficient at infinite dilution and is calculated according to the Othmer-Thakar equation for lysozyme

$$D^{\infty}_{\text{Protein-solution}} = 1.4\,10^{-8}\, \mu_{\text{water}}^{-1.1} V_{Protein}^{-0.6} \tag{37}$$

Here, μ is the viscosity of water in cp and $V_{Protein}$ the molecular volume of the protein in cm^3/mol (Othmer & Thakar, 1953).

8.1.4 Crystal growth

The rate of crystal growth can be expressed as the rate of displacement of a given crystal face in the direction perpendicular to the face v_{hkl} or by the overall linear growth rate G. Each crystallographic face usually exhibits a different linear growth rate. The overall growth rate is defined as the time derivative of the radius of the sphere having a volume equal to the average volume of one crystal according to

$$G = \frac{dL}{dt} \tag{38}$$

Two processes exert a decisive influence during the growth of crystals from solution: mass transport from the solution to the crystal surface by volume diffusion or convection or combination of both mechanisms and incorporation of material into the crystal lattice through the surface integration process (surface reaction process). In most cases, several mechanisms influence the growth rate. If the several mechanisms can be viewed in parallel then the mechanism resulting in the fastest growth controls the overall growth rate. If the processes take place in series as the case for bulk diffusion followed by the surface reaction then the slowest mechanism controls the overall growth rate.

8.1.5 Growth controlled by integration

Surface integration is the process by which the growth units are incorporated to the crystal surface. The two mechanisms controlling the surface integration are (i) two dimensional nucleation and (ii) screw dislocations. These two mechanisms can take place in parallel. The emergence of screw dislocation on a crystal face results in the formation of a growth step on an otherwise smooth crystal surface. If the growth is controlled by diffusion of growth units across the face towards steps on the surface, the growth rate of face is:

$$G_{SD} = \frac{K_{SD1}}{\mathrm{K_{SD2}}} (\mathrm{S}-1)^2 \tanh\left(\frac{\mathrm{K_{SD2}}}{(\mathrm{S}-1)}\right) \tag{39}$$

where K_{SD1} is a function of temperature, retardation factor during adsorption of the growth unit into a kink site in the step, shape of the spiral, number of growth units in unit volume of solution and activation energy of dehydration and K_{SD2} a function of the surface energy, temperature, diameter of the growth units, the number of cooperating spirals. This equation was developed by Burton, Cabrera and Frank (Chernov, 2004).

8.1.6 Supersaturation

The supersaturation is classically written by the difference of solute chemical potential in liquid and solid phase:

$$\frac{\Delta\mu_i}{RT} = \ln\left(\frac{a_i}{a_{i,eq}}\right) = \ln\left(\frac{\gamma_i{}^L C_i}{\gamma_{i,eq}{}^L C_{i,eq}}\right) \tag{40}$$

The supersaturation ratio is defined as

$$S = \frac{\gamma_i{}^L C_i}{\gamma_{i,eq}{}^L C_{i,eq}} \tag{41}$$

A frequent assumption is that $\frac{\gamma_i^{\,L}}{\gamma_{i,eq}^{\,L}} \to 1$ so that the supersaturation ratio can be reduced to the concentration ratios as

$$S = \frac{C_i}{C_{i,eq}} \tag{42}$$

The activity coefficient of a protein can be expressed in terms of the virial coefficients as

$$\ln \gamma_i^{\,L} = 2B_{22}C'_i + \frac{3}{2}B_{222}C'^{\,2}_i + 0(B_{2222}C'^{\,3}_i) \approx 2B_{22}C'_i \tag{43}$$

where the standard state for the protein is taken as $\gamma_i^{\,L} \to 1$ and $C'_i \to 0$. B_{22}, B_{222} and B_{2222} are respectively the second, third and forth virial coefficients and C' the molar concentration (Grant, 2000). In dilute solutions, binary interactions are much more probable than ternary or quaternary interactions so that it is justifiable only to consider pair wise interactions. The combination of equations (13) and (16) gives an estimate for the thermodynamic driving force for crystallisation:

$$\frac{\Delta\mu_i}{k_B T} = \ln\left(\frac{C_i}{C_{i,eq}}\right) + 2B_{22}M_i(C_i - C_{i,eq})\,10^{-3} \tag{44}$$

with C_i as the protein concentration in g protein per kg water.

8.2 Modelling

From the concentration profile of CP01 and CP02, the four parameters K_{SD1}, K_{SD2}, B_{22} and γ have been fitted by minimising following function

$$f = \sum_{i=1}^{N\exp} \sum_{j=1}^{N\text{point}} \frac{\left\{C_{experimental}(i,j) - C_{calculated}(i,j)\right\}^2}{variance(i,j)\; C_{experimental}(i,j)^2} + \sum_{i=1}^{N\exp} \frac{\left\{N_{experimental}(i) - N_{calculated}(i)\right\}^2}{N_{experimental}(i)^2}$$

using the mean value of number crystal concentration, $N_{experimental}$(CP01) = $1.39x10^{11}$/kg and $N_{experimental}$(CP02) = $1.34x10^{10}$/kg taken from later stages of the crystallisation process. An equilibrium concentration of 7.71 mg/mL was assumed. Moreover, the weight of the medium, i.e. the weight of the volume in which the crystallisation process was conducted in, was taken as 1 g, the protein density as 1360 mg/mL, the molecular weight of BHA as 55 kDa and the diameter of the BHA monomer as 6 nm.

8.3 Results and discussion

The concentration profiles and crystal concentrations of CP01 and CP02 were used to calculate the fitting parameter f, the two Burton-Cabrera-Frank growth rate parameters K_{SD1} and K_{SD2}, the second osmotic virial coefficient B_{22} and the surface energy between crystal and solution. Following values have been found:

$f = 0.53$
$K_{SD1} = 1.0778x10^{6}$ m/h
$K_{SD2} = 19.455$
$B_{22} = 0.286x10^{-4}$ mol mL g^{-2}
$\gamma = 0.513$ mJ/m^2

The fitting parameter f is a measure how similar the chosen description of the considered crystallisation processes is. Values <1 generally indicate high similarity whereas values >1 show that the considered processes are probably too different from each other and the predictions of the model will probably not be in good agreement with the experimental data. The model was optimised towards the prediction of the crystal number. The resulting mean number diameter can be used to evaluate the accuracy of the assumptions on which the model is based on.

The found value for the fitting parameter f of 0.53 can be considered as good, meaning that the development of the chosen units leads to a description of the two considered crystallisation processes which suggests a fairly high similarity. The obtained second osmotic

virial coefficient of 0.286×10^{-4} mol mL g^{-2} is in very good agreement with experimental results obtained from self-interaction chromatography for the used conditions (Chapter 6).

Since the model was optimised towards the prediction of the crystal concentration, it is not surprising that the absolute values are in good agreement with the experimental data, both for CP01 and CP02 (Figure 8.1 and Figure 8.2). The model suggests a very rapid increase in crystal concentration such that the final crystal concentration is already reached within the first hour after induction of the supersaturation. This increase is slightly lower in CP02. The experimental increase in crystal concentration is much slower. Interestingly, the increase was slower for CP02 compared to CP01 which is in agreement with the suggestions of the model. As discussed in Chapter 3, the used measuring technique could not detect crystals smaller than 2 µm. As a consequence, a certain time is needed until the crystals exceed the detectable size which could explain why the increase in crystal concentration experimentally observed is slower than the one predicted by the model. In contrast, the prediction of the mean number diameter (7.5 µm for CP01 and 15.6 µm for CP02) differs from the one experimentally determined (4.2 µm for CP01 and 4.7 µm for CP 02). However, the tendencies predicted by the model are correct. Furthermore, the experimentally obtained mean number diameter already reached its final size very shortly after the initiation of the crystallisation process, i.e. already after 20 minutes. The model predicts that the final (but larger) size should almost be reached after five to six hours. In contrast, if the experimentally determined mean number diameter is taken as a measure, the model would also predict that this size should be reached after approximately one hour. Again, the measuring technique may be responsible for the discrepancies. As crystals smaller than 2 µm are not considered, the calculated mean diameter may be higher than actually the case. Had it been possible to consider these particles, a gradual increase in mean diameter would be expected, although only in the very early stages, as also suggested by the model.

Following things should be considered when the difference between the predicted and the measured final crystal size is discussed. The model is based on the closure of the mass balances meaning that lower crystal sizes as experimentally observed would not correlate with the protein concentration profile at the assumed crystal concentration. The crystal concentrations at the end of the crystallisation processes are included in the model which is believed to be appropriate as essentially all crystals should have exceeded the detectable size. It is probably the mean number diameter which may not have been correctly reflected by the

used measuring technique. Reasons could be that the conversion from volume to diameter which assumes spherical crystal shape was inaccurate and may be one reason for the discrepancy between experimental and calculated crystal size, at least for CP01. The differences encountered in CP02 are believed to be far too big as to be ascribed to conversion

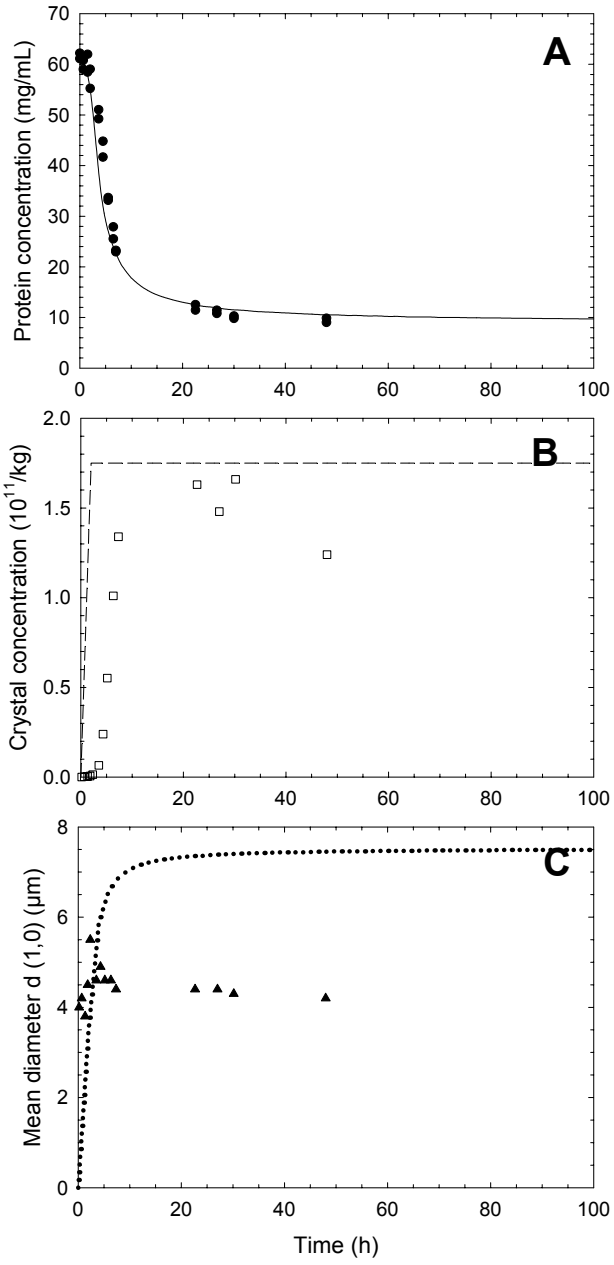

Figure 8.1 Crystallisation process 01. Comparison between experimental results and predictions provided by the model. (A): Concentration profile. (●): Experimental data, (solid line): fitted concentration profile used for the identification of K_{SD1}, K_{SD2}, B_{22} and γ. (B): Crystal concentration. (□) Experimental data, (dashed line): crystal concentration predicted by the model. (C): Mean number diameter d (1,0). (▲): Experimental data, (dotted line): mean diameter predicted by the model.

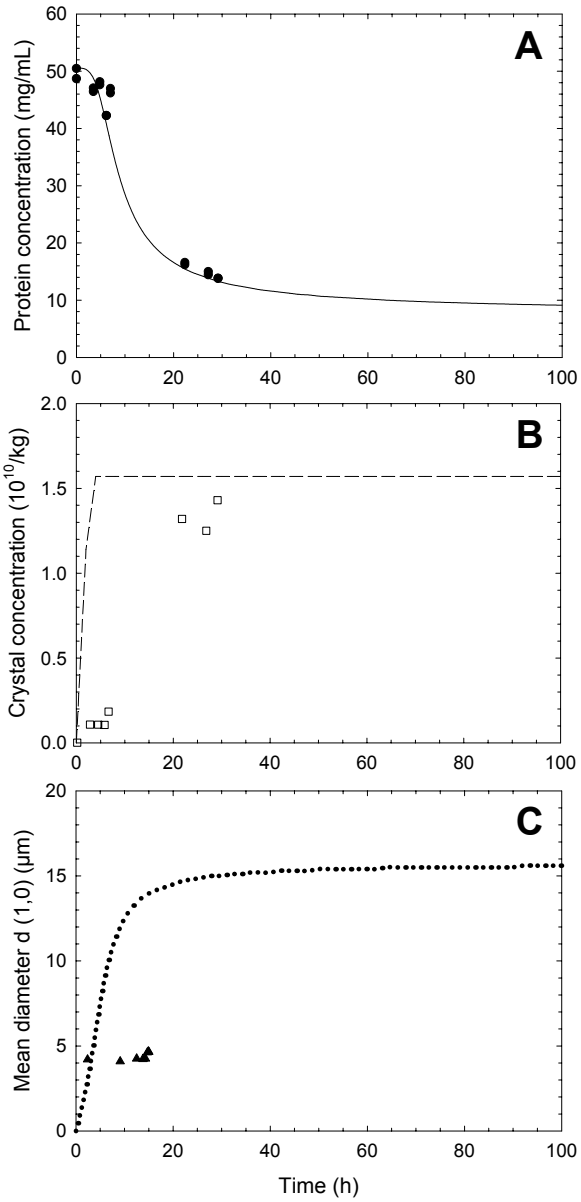

Figure 8.2 Crystallisation process 02. Comparison between experimental results and predictions provided by the model. (A): Concentration profile. (●): Experimental data, (solid line): fitted concentration profile used for the identification of K_{SD1}, K_{SD2}, B_{22} and γ. (B): Crystal concentration. (□) Experimental data, (dashed line): crystal concentration predicted by the model. (C): Mean number diameter d (1,0). (▲): Experimental data, (dotted line): mean diameter predicted by the model.

mistakes from volume to diameter. In this process, the crystal concentration is of approximately one order of magnitude lower than in CP01 whereas the initial supersaturation only differs by 10 mg/mL. As discussed in Chapter 3, CP02 had not yet reached equilibrium and a further increase in crystal concentration towards values observed in CP01 could be expected the more equilibrium is approached.

8.4 Conclusions

The concentration profiles and the crystal numbers at the later stages of two crystallisation processes started from initial supersaturation ratios of 5.5 and 6.6 have been incorporated into a model based on classical nucleation and growth theories. The model was optimised to correctly predict the crystal concentration. The resulting crystal concentrations were in good agreement with the experimental data. However, the model suggested that the final crystal concentrations should already be reached within the first hour of the process whereas a much longer increase was seen experimentally. This discrepancy was ascribed to an incomplete capture of all crystals present in the solution since the used measurement technique had a lower detection limit of 2 μm which may have caused inaccuracies particularly in early phases of the crystallisation process. The calculated final crystal diameter was too high in both crystallisation processes, the deviation was much more pronounced for CP02. The discrepancies were attributed to inaccuracies when the crystal volume was converted into crystal diameters assuming spherical shape which may have been a too gross approximation. The pronounced deviation between experimental and theoretical results as seen in CP02 could have been caused by the fact that the process had not sufficiently reached equilibrium.

In spite of the encountered discrepancies between model and experimental results, the approach taken in this study provides a quantitative inside into the crystallisation kinetics of BHA. Influences e.g., of the initial supersaturation ratio on the development of the supersaturation and the expected mean diameter can be predicted and thus the crystallisation process optimised towards the desired goals. In light of this, the developed model has the potential to be a helpful tool in the development of optimal crystallisation strategies for any given protein.

9 References

1. Agena, S., Pusey, M. L., & Bogle, I. D. (1999); *Biotechnology and Bioengineering,* **64**, 144-150.

2. Anand, K., Pal, D., & Hilgenfeld, R. (2002); *Acta Crystallographica,* **D58**, 1722-1728.

3. Arakawa, T. & Timasheff, S. N. (1985); *Methods in Enzymology,* **114**, 49-77.

4. Ataka, M. & Tanaka, S. (1986); *Biopolymers,* **25**, 337-350.

5. Basu, S. K., Govardhan, C. P., Jung, C. W., & Margolin, A. L. (2004); *Expert opinion on biological therapy,* **4**, 301-317.

6. Behlke, J. & Ristau, O. (1999); *Biophysical Chemistry,* **76**, 13-23.

7. Bénas, P., Legrand, L., & Riès-Kautt, M. (2002); *Acta Crystallographica,* **D58**, 1582-1587.

8. Berger, B. W., Blaney, C. J., Naik, U. P., Bahnson, B. J., & Lenhoff, A. M. (2005); *Crystal Growth and Design,* **5**, 1499-1507.

9. Bergfors, T. (2003); *Journal of Structural Biology,* **142**, 66-76.

10. Betts, L., Frick, L., Wolfenden, R., & Carter, C. W., Jr. (1989); *Journal of Biological Chemistry,* **264**, 6737-6740.

11. Bisgaard-Frantzen, H., Svendsen, A., Norman, B., Pedersen, S., Kjaerulff, S., Outtrup, H., & Borchert, T. V. (1999); *Journal of Applied Glycoscience,* **46**, 199-206.

12. Boistelle, R. & Astier, J. P. (1992); *Joural of Crystal Growth,* **123**, 109-120.

13. Boistelle, R. & Astier, J. P. (1988); *Journal of Crystal Growth,* **90**, 14-30.

14. Boström, M., Williams, D. R. M., & Ninham, B. W. (2003); *Biophysical Journal,* **85**, 686-694.

15. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., & Snel, L. (1988); *Nature,* **333**, 679-682.

16. Broide, M. L., Tominc, T. M., & Saxowsky, M. D. (1996); *Physical Review E,* **53**, 6325-6335.

17. Brzozowski, A. M., Lawson, B. M., Turkenburg, J. P., Bisgaard-Frantzen, H., Svendsen, A., Borchert, T. V., Dauter, Z., Wilson, K. S., & Davies, G. J. (2000); *Biochemistry,* **39**, 9099-9107.

18. Bunn, C. W., Moews, P. C., & Baumber, M. E. (1971); *Proceedings of the Royal Society of London Series B,* **178**, 245-258.

19. Buque-Taboada, E., Straathof, A., Heijnen, J., & van der Wielen, L. A. M. (2004); *Biotechnology and Bioengineering,* **86**, 795-800.

20. Cacioppo, E. & Pusey, M. L. (1991a); *Journal of Crystal Growth,* **110**, 66-71.

21. Cacioppo, E. & Pusey, M. L. (1991b); *Journal of Crystal Growth,* **114**, 286-292.

22. Carbonnaux, C., Riès-Kautt, M., & Ducruix, A. (1995); *Protein Science,* **4**, 2123-2128.

23. Carlsen, M. & Nielsen, J. (2001); *Applied Microbiology and Biotechnology,* **57**, 346-349.

24. Carter, C. W., Jr. & Carter, C. W. (1979); *Journal of Biological Chemistry,* **254**, 12219-12223.

25. Chayen, N. E. (2005); *Progress in Biophysics and Molecular Biology,* **88**, 329-337.

26. Chernov, A. A. (1997); *Physics Reports,* **288**, 61-75.

27. Chernov, A. A. (2004); *Journal of Crystal Growth,* **264**, 499-518.

28. Chernov, A. A. (2003); *Journal of Structural Biology,* **142**, 3-21.

29. Collins, K. D. & Washabaugh, M. W. (1985); *Quarterly Review of Biophysics,* **18**, 323-422.

30. D'Arcy, A. (1993); *Acta Crystallographica,* **D50**, 469-471.

31. Dai, G. L., Yu, Y., Kang, Q., & Hu, W. R. (2004); *Acta Chimica Sinica,* **62**, 757-761.

32. Davies, G. J., Brzozowski, A. M., Dauter, Z., Rasmussen, M. D., Borchert, T. V., & Wilson, K. S. (2005); *Acta Crystallographica,* **D61**, 190-193.

33. Debye, P. & Hückel, E. (1923); *Physikalische Zeitschrift,* **9**, 185-207.

34. Declerck, N., Joyet, P., Trosset, J. Y., Garnier, J., & Gaillardin, C. (1995); *Protein Engineering,* **8**, 1029-1037.

35. Declerck, N., Machius, M., Chambert, R., Wiegand, G., & Gaillardin, C. (1997); *Protein Engineering,* **10**, 541-549.

36. Declerck, N., Machius, M., Wiegand, G., Huber, R., & Gaillardin, C. (2004); *Journal of molecular Biology,* **301**, 1041-1057.

37. DeLucas, L. J., Hamrick, D., Cosenza, L., Nagy, L., McCombs, D., Bray, T., Chait, A., Stoops, B., Belgovskiy, A., & Wilson, W. W. (2005); *Progress in Biophysics and Molecular Biology,* **88**, 285-309.

38. Demoruelle, K., Guo, B., Kao, S., McDonald, H., Nikic, D., Holman, S. C., & Wilson, W. W. (2002); *Acta Crystallographica,* **D58**, 1544-1548.

39. DePhillips, P. & Lenhoff, A. M. (2000); *Journal of Chromatography A,* **883**, 39-54.

40. dos Santos, L. F., Defrenne, L., & Krebs-Brown, A. (2002); *Analytica Chimica Acta,* **456**, 41-54.

41. Ducruix, A., Guilloteau, J. P., Ries-Kautt, M. M., & Tardieu, A. (1996); *Journal of Crystal Growth,* **168**, 28-39.

42. Ducruix, A. F. & Ries-Kautt, M. M. (1990); *A Companion to Methods in Enzymology,* **1**, 25-30.

43. Durbin, S. D. & Feher, G. (1986); *Journal of Crystal Growth,* **76**, 583-592.

44. Edsall, J. T. (1930); *Journal of Biological Chemistry,* **89**, 289-313.

45. El Bawab, S., Bielawska, A., & Hannun, Y. A. (1999); *Journal of Biological Chemistry,* **39**, 27948-27955.

46. Feher, G. (1986); *Journal of Crystal Growth,* **76**, 545-546.

47. Feher, G. & Kam, Z. (1985); *Methods in Enzymology,* **114**, 77-112.

48. Feigelson, R. S. (1988); *Journal of Crystal Growth,* **90**, 1-13.

49. Fish, N. M. & Lilly, M. D. (1984); *Nature Biotechnology,* **2**, 623-627.

50. Forsythe, E. & Pusey, M. L. (1994); *Journal of Crystal Growth,* **139**, 89-94.

51. Fritsch, G., Koepke, J., Diem, R., Kuglstatter, A., & Baciou, L. (2002); *Acta Crystallographica,* **D58**, 1660-1663.

52. García, C. D., Hadley, D. J., Wilson, W. W., & Henry, C. S. (2003); *Biotechnology Progress,* **19**, 1006-1010.

53. García-Ruiz, J. M. (2003); *Journal of Structural Biology,* **142**, 22-31.

54. Georgalis, Y. & Saenger, W. (1999); *Science Progress,* **82**, 271-294.

55. George, A. & Wilson, W. W. (1994); *Acta Crystallographica,* **D50**, 361-365.

56. Gernert, K. M., Smith, R., & Carter, D. C. (1988); *Analytical Biochemistry,* **168**, 141-147.

57. Giegé, R., Dock, A. C., Kern, D., Lorber, B., Thierry, J. C., & Moras, D. (1986); *Journal of Crystal Growth,* **76**, 554-563.

58. Giegé, R., Drenth, J., Ducruix, A., McPherson, A., & Saenger, W. (1995); *Progress in Crystal Growth and Characterization,* **30**, 237-281.

59. Grant, M. L. (2000); *Journal of Crystal Growth,* **209**, 130-137.

60. Green, A. A. (1932); *Physical Chemistry of Proteins,* **10**, 47-66.

61. Guilloteau, J. P., Riès-Kautt, M., & Ducruix, A. (1992); *Journal of Crystal Growth,* **122**, 223-230.

62. Guo, B., Kao, S., McDonald, H., Asanov, A., Combs, L., & Wilson, W. W. (1999); *Journal of Crystal Growth,* **196**, 424-433.

63. Haas, C., Drenth, J., & Wilson, W. W. (1999); *Journal of Physical Chemistry B,* **103**, 2808-2811.

64. Hanko, V. P. & Rohrer, J. S. (2004); *Analytical Biochemistry,* **324**, 29-38.

65. Haynes, C. A., Tamura, K., Korfer, H. R., Blanch, H. W., & Prausnitz, J. M. (1992); *Journal of Physical Chemistry,* **96**, 905-912.

66. He, X. X., Wang, K. M., Tan, W. H., Lin, X., Chen, L., & Chen, X. H. (2003); *Reviews on Advanced Materials Science,* **5**, 375-380.

67. Henrissat, B. (1991); *Biochemical Journal,* **280**, 309-316.

68. Hirsch, R. E., Lin, M. J., & Nagel, R. L. (1988); *Journal of Biological Chemistry,* **263**, 5936-5939.

69. Hofmeister, F. (1888); *Archiv für experimentelle Pathologie und Pharmakologie,* **24**, 247-260.

70. Jacobsen, C. (1998); *PhD-thesis, University College London,* UK.

71. Jacobsen, C., Garside, J., & Hoare, M. (1998); *Biotechnology and Bioengineering,* **57**, 666-675.

72. Jelesarov, I., Dürr, E., Thomas, R. M., & Bosshard, H. R. (1998); *Biochemistry,* **37**, 7539-7550.

73. Jones, R. M. (2003); *American Laboratory,* **1**, 44-47.

74. Judge, R. A., Johns, M. R., & White, E. T. (1995); *Biotechnology and Bioengineering,* **48**, 316-323.

75. Judge, R. A., Johns, M. R., & White, E. T. (1996); *Journal of Chemical and Engineering Data,* **41**, 422-424.

76. Kadima, W., Ogendal, L., Bauer, R., Kaarsholm, N., Brodersen, K., Hansen, J. F., & Porting, P. (1993); *Biopolymers,* **33**, 1643-1657.

77. Kam, Z., Shore, H. B., & Feher, G. (1978); *Journal of molecular Biology,* **123**, 539-555.

78. Kashchiev, D. & van Rosmalen, G. M. (2003); *Crystal Research and Technology,* **38**, 555-574.

79. Kiesslich, T., Oberdanner, C., Krammer, B., & Plaetzer, K. (2003); *Journal of Biochemical and Biophysical Methods,* **57**, 247-251.

80. Kirkwood, J. G. (1934); *Journal of Chemical Physics,* **2**, 351-361.

81. Klyushnichenko, V. (2003); *Current opinion in drug discovery & development,* **6**, 848-854.

82. Kreusch, S., Schwedler, S., Tautkus, B., Cumme, G. A., & Horn, A. (2003); *Analytical Biochemistry,* **313**, 208-215.

83. Krieger, N., Taipa, M. A., Melo, E. H. M., Lima-Filho, J. L., Aires-Barros, M. R., & Cabral, J. M. S. (1999); *Bioprocess and Biosystems Engineering,* **20**, 59-65.

84. Leavis, P. C. & Rothstein, F. (1974); *Archives of Biochemistry and Biophysics,* **161**, 671-682.

85. Lee, H. M., Kim, Y. W., & Baird, J. K. (2001); *Journal of Crystal Growth,* **232**, 294-300.

86. Lee, T. S., Vaghjiani, J. D., Lye, G. J., & Turner, M. K. (2000); *Enzyme and Microbial Technology,* **26**, 582-592.

87. Lin, D. Q., Brixius, P. J., Hubbuch, J. J., Thömmes, J., & Kula, M. R. (2003); *Biotechnology and Bioengineering,* **83**, 149-157.

88. Lorentz, K. (2000); *Clinical Chemistry,* **46**, 644-649.

89. Lyhne-Iversen, L. (2005); *Master thesis, Technical University of Denmark,* Denmark.

90. Machius, M., Declerck, N., Huber, R., & Wiegand, G. (1998); *Structure,* **6**, 281-292.

91. Machius, M., Wiegand, G., & Huber, R. (1995); *Journal of molecular Biology,* **246**, 545-559.

92. Malkin, A. J., Kuznetsov, Y. G., Glantz, W., & McPherson, A. (1996); *Journal of Physical Chemistry,* **100**, 11736-11743.

93. Margolin, A. L. (1996); *Trends in Biotechnology,* **14**, 223-230.

94. Matsushita, M., Irino, T., Komoda, T., & Sakagishi, Y. (1993); *Clinica Chimica Acta,* **216**, 103-111.

95. Mattison, K. W. & Kaszuba, M. (2004); *American Biotechnology Laboratory,* **12**, 8-11.

96. McDonald, A. (1998); *Materials World,* **6**, 399-401.

97. McPherson, A., Malkin, A. J., & Kuznetsov, Y. G. (1995); *Structure,* **3**, 759-768.

98. Melander, W. & Horvath, C. (1977); *Archives of Biochemistry and Biophysics,* **183**, 200-215.

99. Moffett, J. R., Namboodiri, M. A., & Neale, J. H. (1993); *Journal of Histochemistry and Cytochemistry,* **41**, 559-570.

100. Mullet, M., Fievet, P., Reggiani, J. C., & Pagetti, J. (1997); *Journal of Membrane Science,* **123**, 255-265.

101. Neal, B. L., Asthagiri, D., & Lenhoff, A. M. (1998); *Biophysical Journal,* **75**, 2469-2477.

102. Neal, B. L. & Lenhoff, A. M. (1995); *Annals of the International Chemical Engineering Journal,* **41**, 1010-1014.

103. Ng, J. D., Lorber, B., Witz, J., Théobald-Dietrich, A., Kern, D., & Giegé, R. (1996); *Journal of Crystal Growth,* **168**, 50-62.

104. Nielsen, A. D. (2003); *PhD-thesis, Roskilde University,* Denmark.

105. Nielsen, A. D., Fuglsang, C. C., & Westh, P. (2003); *Biochemical Journal,* **373**, 337-343.

106. Nielsen, J. E. & Borchert, T. V. (2000); *Biochimica et Biophysica Acta,* **1543**, 253-274.

107. Ninomiya, K., Yamamoto, T., Oheda, T., Sato, K., Sazaki, G., & Matsuura, Y. (2001); *Journal of Crystal Growth,* **222**, 311-316.

108. Olsen, S. & Falholt, P. (1998); *Journal of Surfactants and Detergents,* **1**, 555-567.

109. Othmer, D. F. & Thakar, M. S. (1953); *Industrial and Engineering Chemistry,* **45**, 589-593.

110. Pan, X. & Glatz, C. E. (2003); *Crystal Growth and Design,* **3**, 203-207.

111. Patro, Y. P. & Przybycien, T. M. (1996); *Biotechnology and Bioengineering,* **52**, 193-203.

112. Payan, F. & Qian, M. (2003); *Journal of Protein Chemistry,* **22**, 275-284.

113. Pearson, C. R., Heng, M., Gebert, M., & Glatz, C. E. (2004); *Biotechnology and Bioengineering,* **87**, 54-60.

114. Pearson, R. G. (1963); *Journal of the American Chemical Society,* **85**, 3533-3539.

115. Petsev, D. & Vekilov, P. G. (2000); *Physical Review Letters,* **84**, 1339-1342.

116. Petsev, D. N., Thomas, B. R., Yau, S.-T., Tsekova, D., Nanev, C., Wilson, W. W., & Vekilov, P. G. (2001); *Journal of Crystal Growth,* **232**, 21-29.

117. Prouty, W. F. (1991); *Bioprocess Technology,* **13**, 221-262.

118. Pusey, M. L. & Munson, S. (1991); *Journal of Crystal Growth,* **113**, 385-389.

119. Retailleau, P., Ducruix, A., & Riès-Kautt, M. (2002); *Acta Crystallographica,* **D58**, 1576-1581.

120. Retailleau, P., Riès-Kautt, M., & Ducruix, A. (1997); *Biophysical Journal,* **73**, 2156-2163.

121. Ricq, L., Pierre, A., Reggiani, J. C., Pagetti, J., & Foissy, A. (1998); *Colloids and Surfaces A: Physicochemical and Engineering Aspects,* **138**, 301-308.

122. Riès-Kautt, M. & Ducruix, A. (1997); *Methods in Enzymology,* **276**, 23-59.

123. Riès-Kautt, M. & Ducruix, A. (1989); *Journal of Biological Chemistry,* **264**, 745-748.

124. Riès-Kautt, M. & Ducruix, A. (1991); *Journal of Crystal Growth,* **110**, 20-25.

125. Rohani, S., Tavare, N. S., & Garside, J. (1990); *Canadian Journal of Chemical Engineering,* **68**, 261-267.

126. Rosenberger, F., Howard, S. B., Sowers, J. W., & Nyce, T. A. (1993); *Journal of Crystal Growth,* **129**, 1-12.

127. Rosenberger, F., Vekilov, P. G., Muschol, M., & Thomas, B. R. (1996); *Journal of Crystal Growth,* **168**, 1-27.

128. Ruppert, S., Sandler, S. I., & Lenhoff, A. M. (2001); *Biotechnology Progress,* **17**, 182-187.

129. Saridakis, E. & Chayen, N. E. (2003); *Biophysical Journal,* **84**, 1218-1222.

130. Scatchard, G., Scheinberg, H., & Armstrong, S. H. (1950); *American Chemical Society,* **72**, 540-546.

131. Schmidt, S., Havekost, D., Kaiser, K., Kauling, J., & Henzler, H.-J. (2005); *Engineering in Life Sciences,* **5**, 273-276.

132. Schügerl, K. (2000); *Biotechnology Advances,* **18**, 581-599.

133. Sehgal, D. & Vijay, I. K. (1994); *Analytical Biochemistry,* **218**, 87-91.

134. Sha, Z. L., Hatakka, H., Louhi-Kultanen, M., & Palosaari, S. (1996); *Journal of Crystal Growth,* **166**, 1105-1110.

135. Shi, D., Mhaskar, P., El Farra, N. H., & Christofides, P. D. (2005); *Nanotechnology,* **16**, S562-S574.

136. Skouri, M., Lorber, B., Giege, R., Munch, J.-P., & Candau, J. S. (1995); *Journal of Crystal Growth,* **152**, 209-220.

137. Spalding, B. J. (1991); *Bio/Technology,* **9**, 229-232.

138. Staby, A., Johansen, N., Wahlstrom, H., & Mollerup, I. (1998); *Journal of Chromatography A,* **827**, 311-318.

139. Stirpe, A., Guzzi, R., Verbeet, M. P., Canters, G. W., & Sportelli, L. (2002); *Journal of Inorganic Biochemistry,* **91**, 463-469.

140. Suvd, D., Fujimoto, Z., Takase, K., Matsumura, M., & Mizuno, H. (2001); *Journal of biochemistry,* **129**, 461-468.

141. Synowiec, P., Jones, A. G., & Ayazi Shamlou, P. (1993); *Chemical Engineering Science,* **48**, 3485-3495.

142. Tanford, C. (1961); *Physical Chemistry of Macromolecules*, Wiley, New York, USA.

143. Tardieu, A., Bonnet,, F., Finet, S., & Vivarès, D. (2002); *Acta Crystallographica,* **D58**, 1549-1553.

144. Tavare, N. S. (1986); *Chemical Engineering Communications,* **61**, 259-318.

145. Tavare, N. S. (1991); *Reviews in Chemical Engineering,* **7**, 298-309.

146. Tavares, F. W., Bratko, D., Striolo, A., Blanch, H. W., & Prausnitz, J. M. (2004); *Journal of Chemical Physics,* **120**, 9859-9869.

147. Tessier, P. M., Johnson, H. R., Pazhianur, R., Berger, B. W., Prentice, J. L., Bahnson, B. J., Sandler, S. I., & Lenhoff, A. M. (2003); *Proteins: Structure, Function and Genetics,* **50**, 303-311.

148. Tessier, P. M. & Lenhoff, A. M. (2003); *Current opinion in Biotechnology,* **14**, 512-516.

149. Tessier, P. M., Lenhoff, A. M., & Sandler, S. I. (2002a); *Biophysical Journal,* **82**, 1620-1631.

150. Tessier, P. M., Vandrey, S. D., Berger, B. W., Pazhianur, R., Sandler, S. I., & Lenhoff, A. M. (2002b); *Acta Crystallographica,* **D58**, 1531-1535.

151. Timasheff, S. N. & Arakawa, T. (1988); *Journal of Crystal Growth,* **90**, 39-46.

152. Toyokura, K. (1995); *Journal of Chemical Engineering of Japan,* **28**, 361-371.

153. Upadek, H. & Kottwitz, B. (1992); *Chimica Oggi,* **12**, 203-212.

154. Vallee, B. L., Stein, E. A., Sumerwell, W. N., & Fischer, E. H. (1959); *Journal of Biological Chemistry,* **234**, 2901-2095.

155. van der Maarel, J. J. E. C., van der Veen, B., Uitehaag, J. C. M., Leemhuis, H., & Dijkhuizen, L. (2002); *Journal of Biotechnology,* **94**, 137-155.

156. van Ee, J. H. (1992); *Chimica Oggi,* **10**, 4-20.

157. Vaney, M. C., Broutin, L., Retailleau, P., Douangamath, A., Lafont, S., Hamiaux, C., Prange, T., Ducruix, A., & Riès-Kautt, M. (2001); *Acta Crystallographica,* **D57**, 929-940.

158. Veesler, S., Lafont, S., Marcq, S., Astier, J. P., & Boistelle, R. (1996); *Journal of Crystal Growth,* **168**, 124-129.

159. Veesler, S., Marcq, S., Lafont, S., Astier, J. P., & Boistelle, R. (1993); *Acta Crystallographica,* **D50**, 355-360.

160. Vekilov, P. G., Feeling-Taylor, A. R., Yau, S. T., & Petsev, D. (2002); *Acta Crystallographica,* **D58**, 1611-1616.

161. Vekilov, P. G., Monaco, L. A., Thomas, B. R., Stojanoff, V., & Rosenberger, F. (1996); *Acta Crystallographica,* **D52**, 785-798.

162. Velev, O. D., Kaler, E. W., & Lenhoff, A. M. (1998); *Biophysical Journal,* **75**, 2682-2697.

163. Vilker, V. L., Colton, C. K., & Smith, K. A. (1981); *Journal of Colloid and Interface Science,* **79**, 548-566.

164. Vuolanto, A., Uotila, S., Leisola, M., & Visuri, K. (2003); *Journal of Crystal Growth,* **257**, 403-411.

165. Walsh, G. & Headon, D. (1994); *Protein Biotechnology*, John Wiley & Sons Ltd., England.

166. Weber, P. C. (1991); *Advances in Protein Chemistry,* **41**, 1-35.

167. Weiss, M. S., Palm, G. J., & Hilgenfeld, R. (2000); *Acta Crystallographica,* **D56**, 952-958.

168. Wilson, W. W. (2003); *Journal of Structural Biology,* **142**, 56-65.

169. Windsor, D. (2003); *Analytical Biochemistry,* **325**, 1-20.

170. Wyatt, P. J. (1993); *Analytica Chimica Acta,* **272**, 1-40.

171. Wynn, E. J. W. & Hounslow, M. J. (1997); *Powder Technology,* **93**, 163-175.

172. Zimm, B. (1948); *Journal of Chemical Physics,* **16**, 1093-1099.